Listening to the World's Oceans: Searching for Marine Mammals by Detecting and Classifying Terabytes of Bioacoustic Data in Clouds of Noise

Christopher W. Clark, Peter J. Dugan, Dimitri W. Ponirakis, Marian Popescu, Mohammad Pourhomayoun, Yu Shiu, John Zelweg

Bioacoustics Research Program, Cornell Lab of Ornithology, Cornell University, Ithaca, New York 148504, USA

http://www.birds.cornell.edu/brp/

Acknowledgements

We gratefully recognize the following people for making this work possible:

- Michael Weise and Dana Belden, the Office of Naval Research (ONR grant N000141210585),
- Allison Miller, National Oceanic Partnership Program,
- Will Jackson, the National Fish and Wildlife Foundation (grant 0309.07.28515),
- Angelo D'Amato, the MathWorks, Inc.

We also acknowledge those who providing data, data products and intellectual inspiration for this work:

- Cornell Bioacoustics Research deployment and retrieval team, Research science team, and the Detection-classification, high-performance-computing (HPC) team
- Chuck Gagnon (USN LCR Ret.) – Jedi acoustic tracker
- Leila Hatch, David Wiley and Sofie Van Parijs, NOAA Sanctuaries and NOAA NESFC
- Roger Payne and Katy Payne – whale song
- William T. Ellison, Marine Acoustics, Inc. – underwater acoustics guru

The Grand Illusion

In search of an automated solution for detecting and identifying animal sounds in BIG data sets

“Beam me up, Scottie!”

Three Basic Messages

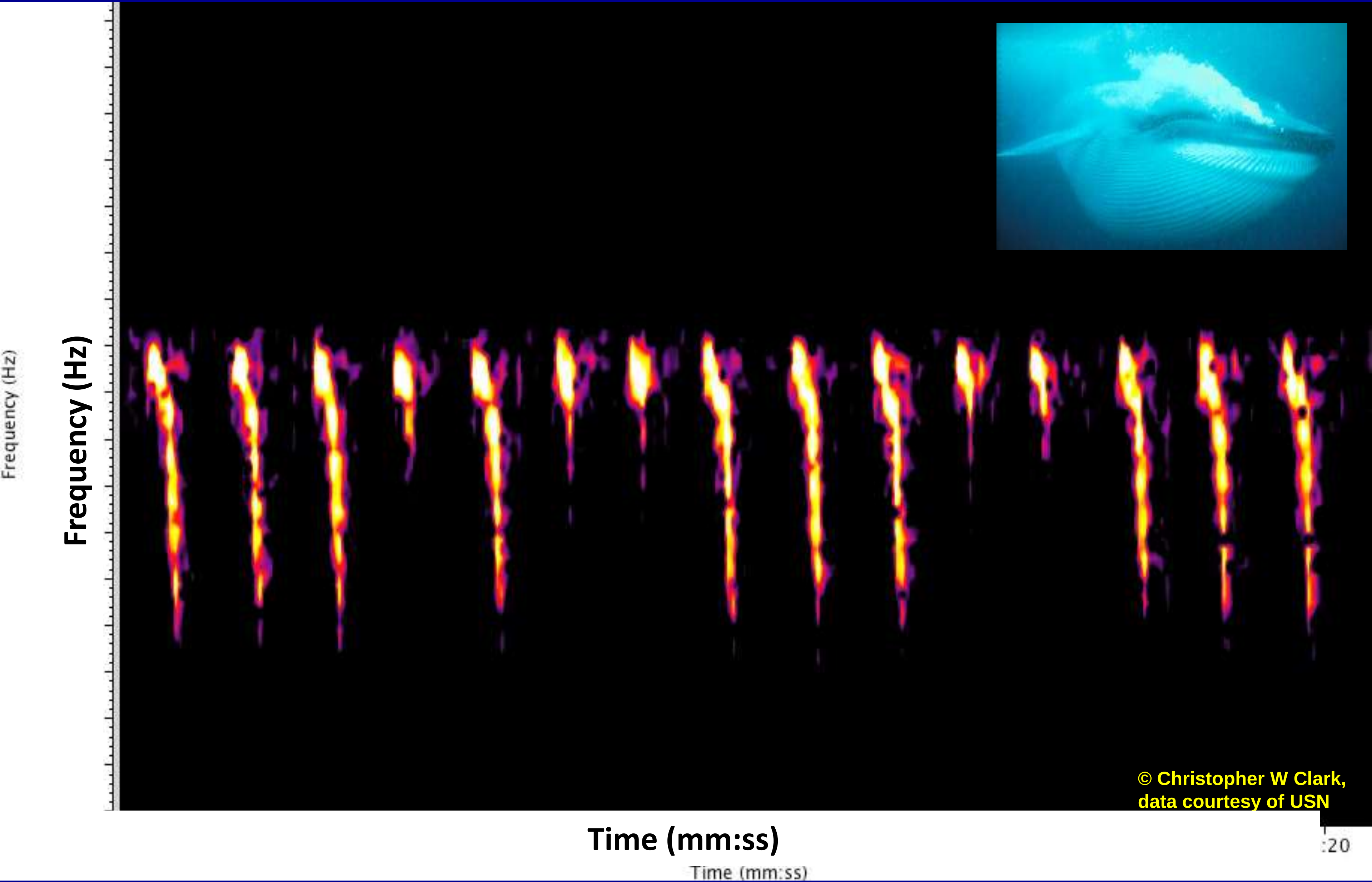
- **The spatio-temporal-spectral scales of the problem: Marine mammals produce a great variety of sounds and depend on sound and their acoustic environments for basic life functions** (*Acoustic ecology*).

- **It is critical to process acoustic data at large scales. Human activities impose huge risks to whales and all marine life over very large spatial and temporal scales.** (*Chronic noise from shipping and offshore energy exploration*).

- **Why synthesis of these data products makes a difference. We must acquire knowledge to change the conceptual paradigm, our attitudes, and our behaviors** (scientific activism)**!**

The Ocean is Alive with the Sounds of Life.

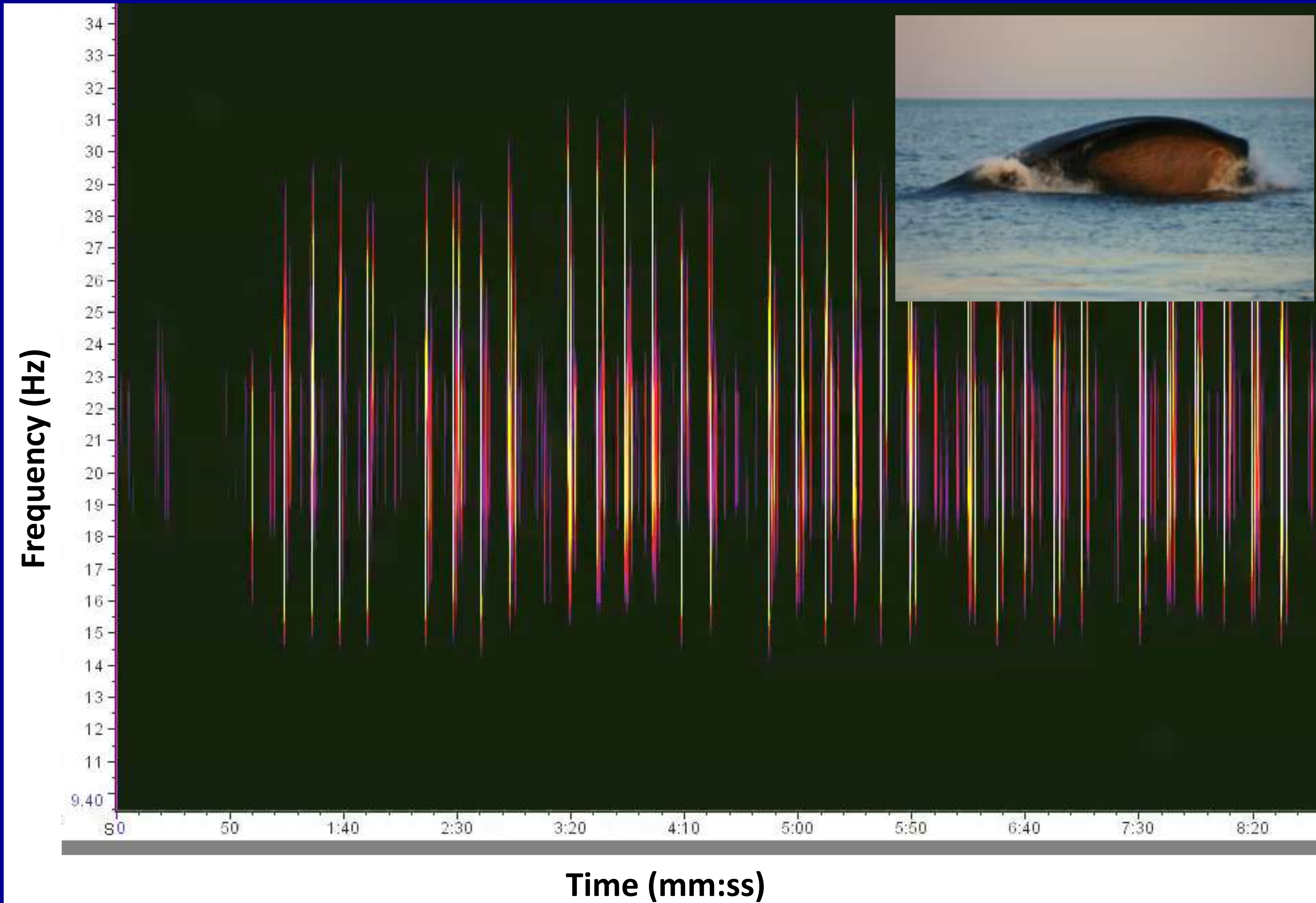
Blue whale singers can be heard across an ocean.
Here at x30; One song note = 15-19 Hz, 20 sec, 2000km



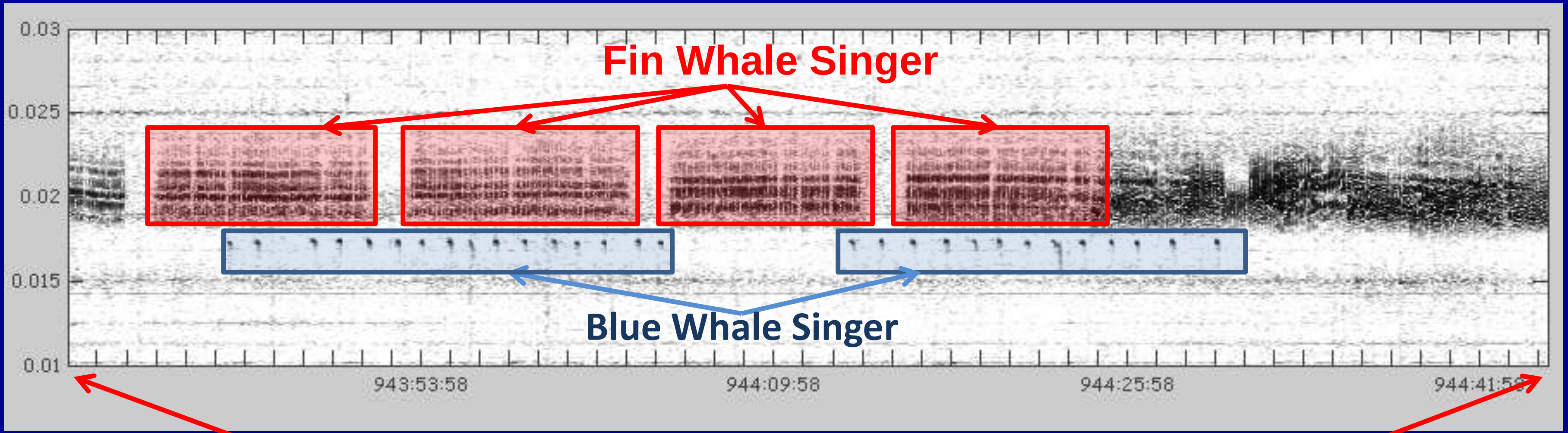
© Christopher W Clark,
data courtesy of USN

Fin Whale at x30 -- Deep Water, Cosmopolitan

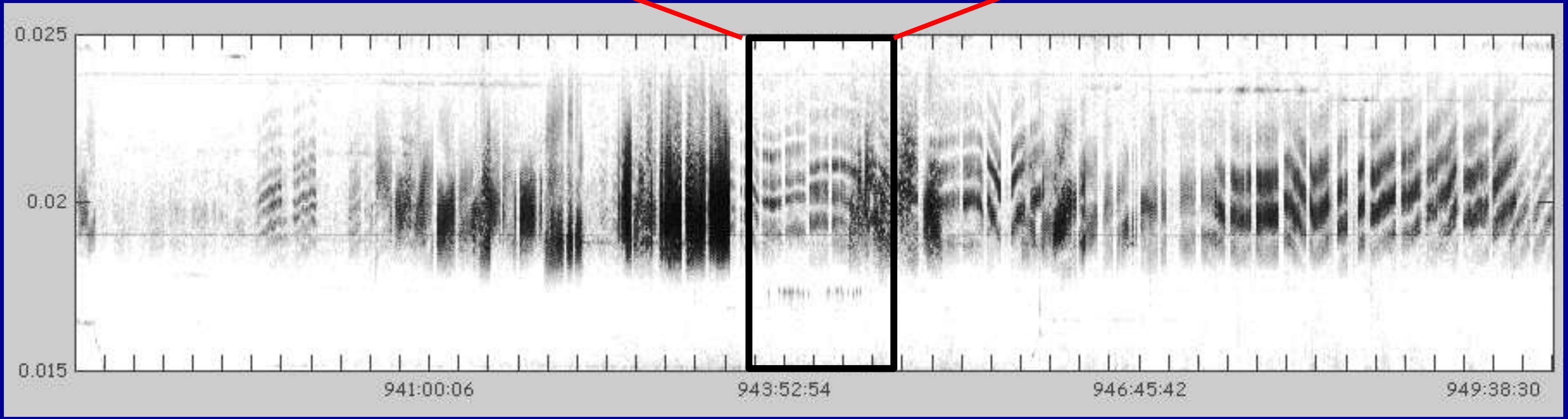
One song note = 18-25 Hz, 1 sec, 1000km



Blue Whale and Fin Whale Songs

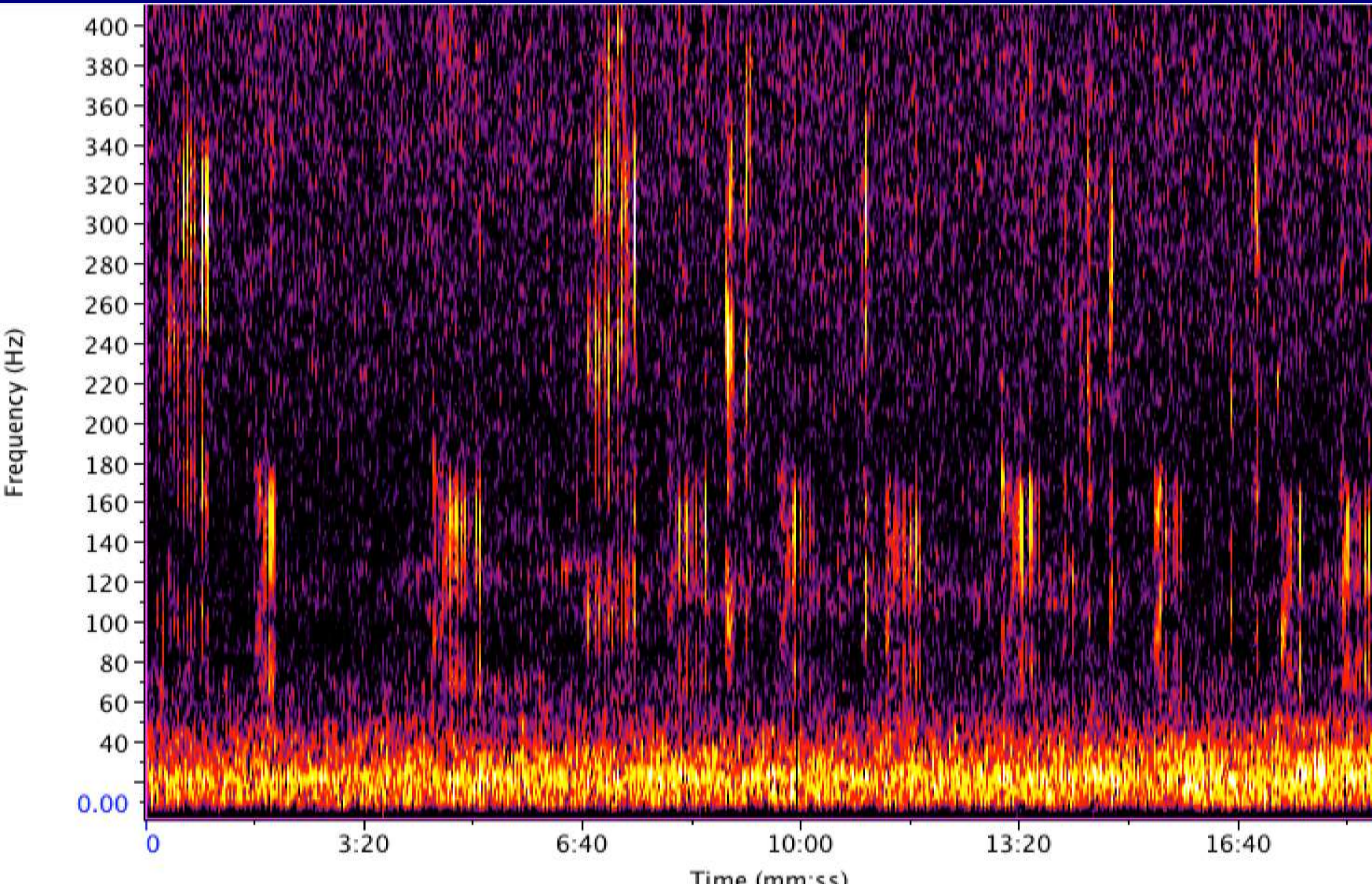


2 hours

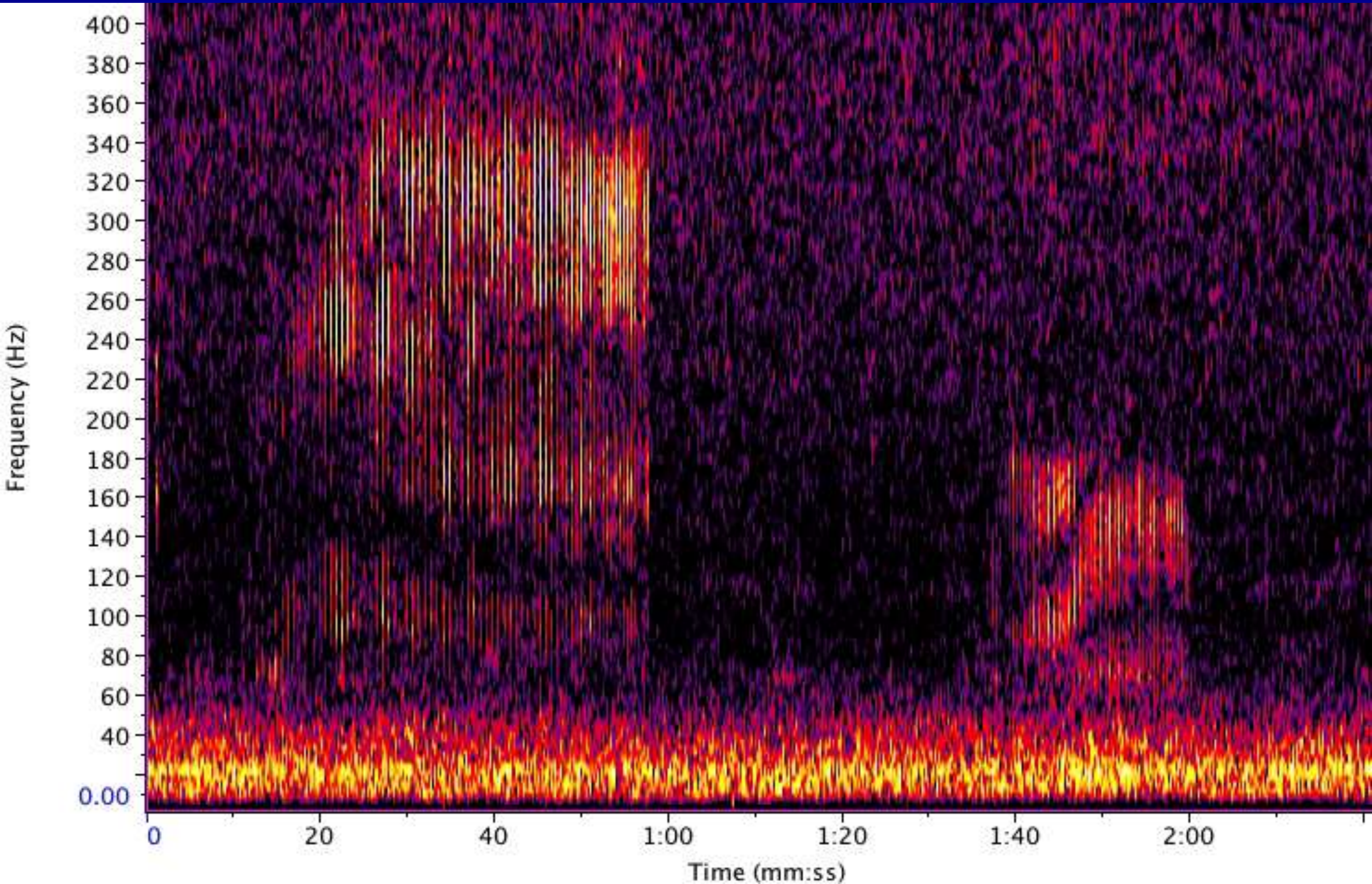


24 hours

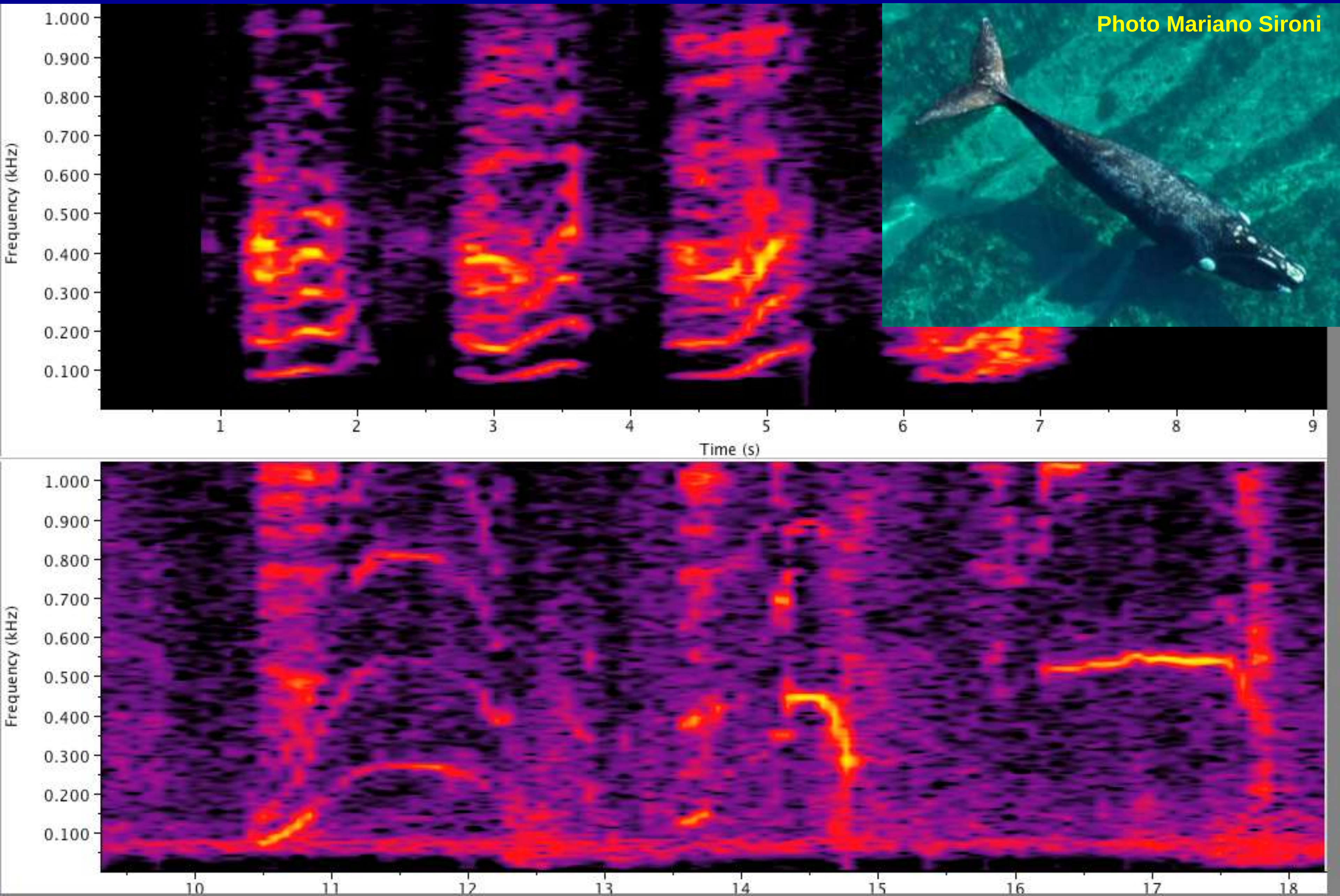
Minke Whales -- Cosmopolitan
One song = 40-1000 Hz, 20-80 sec, 15km



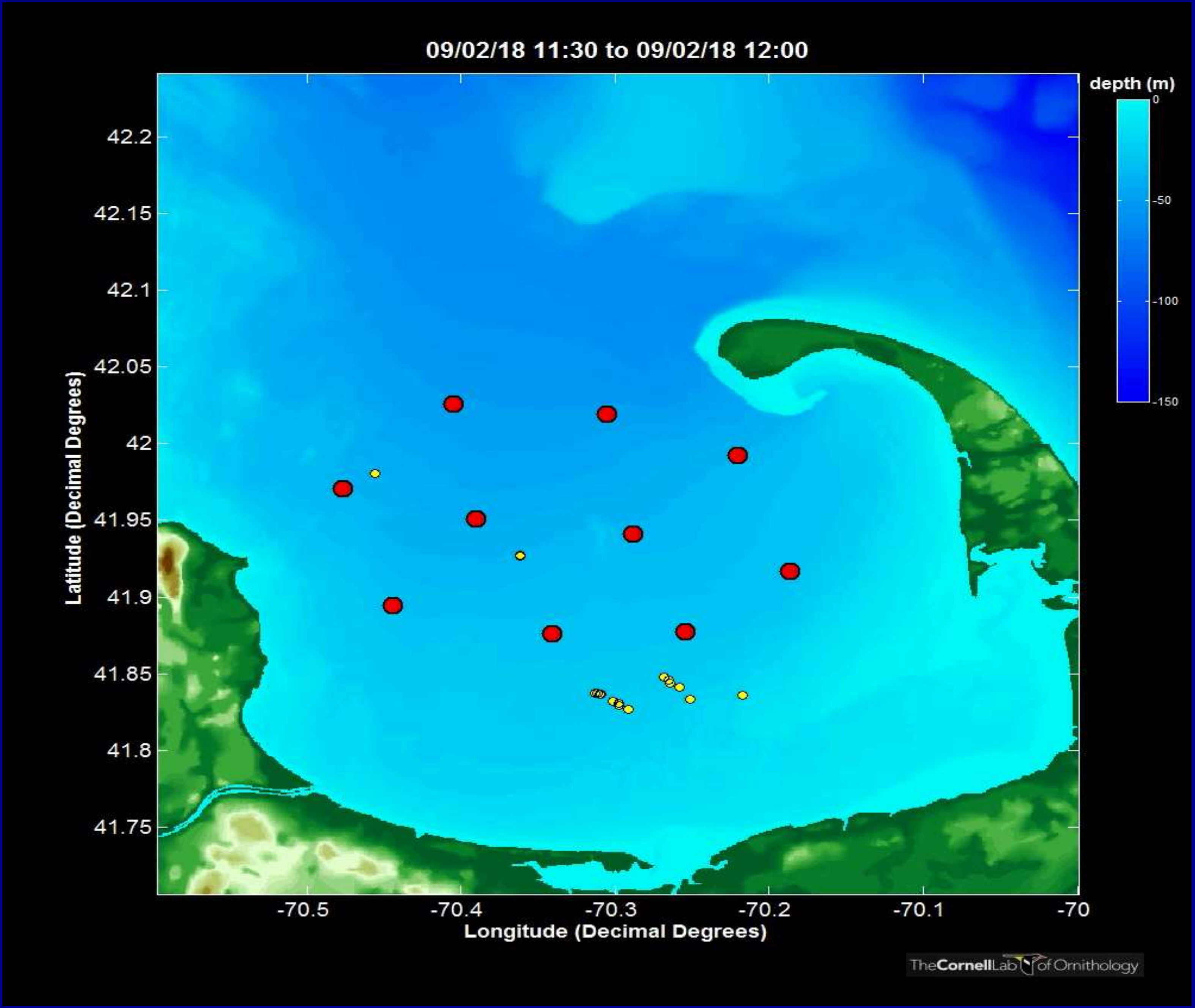
Two Minke Whale Songs
Sped up x15



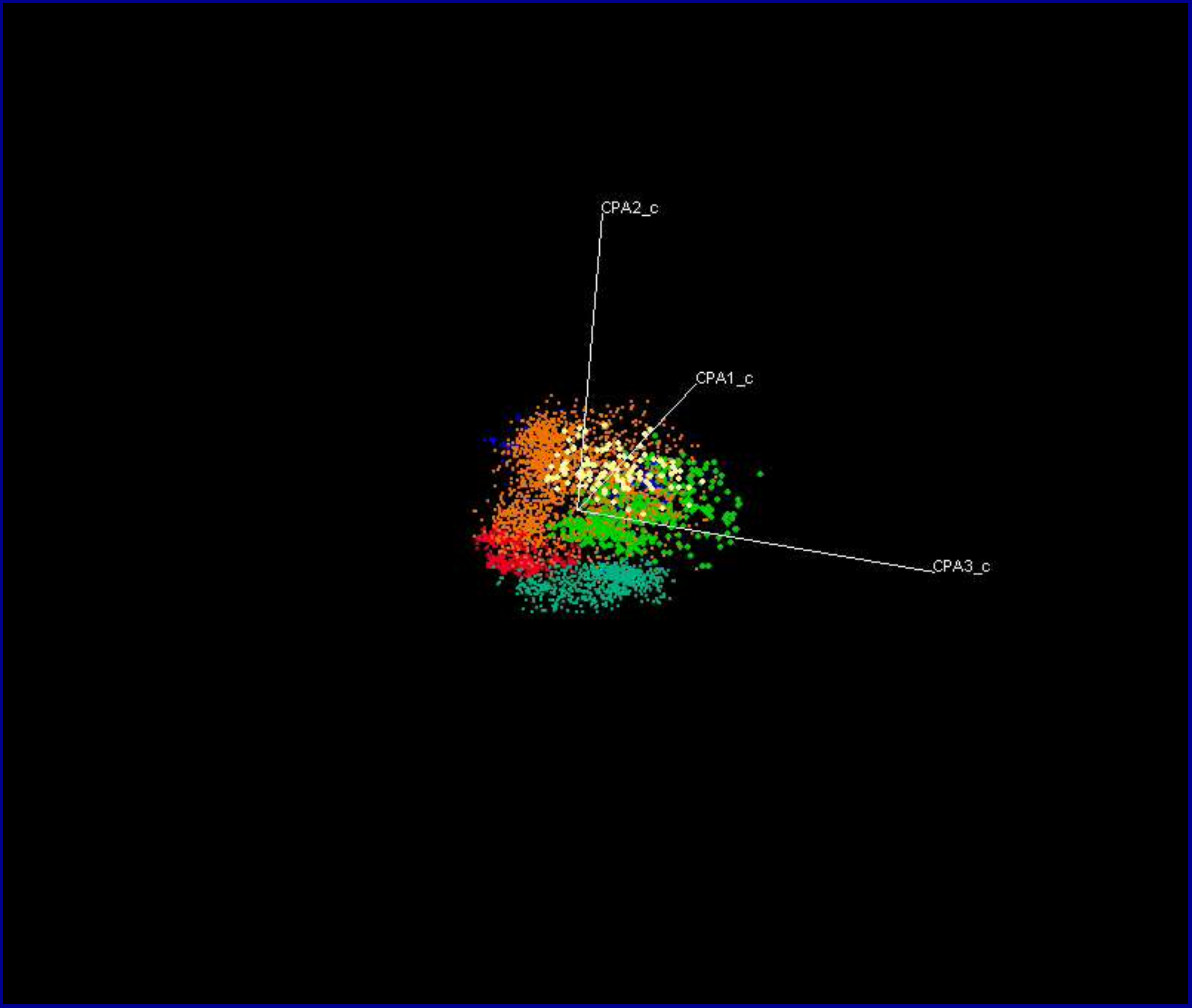
Right Whales – Coastal, Highly Endangered



Right Whale Acoustic Communication: Their Social Network



Bioacoustic Feature Space for Whales



Blue

Bowhead

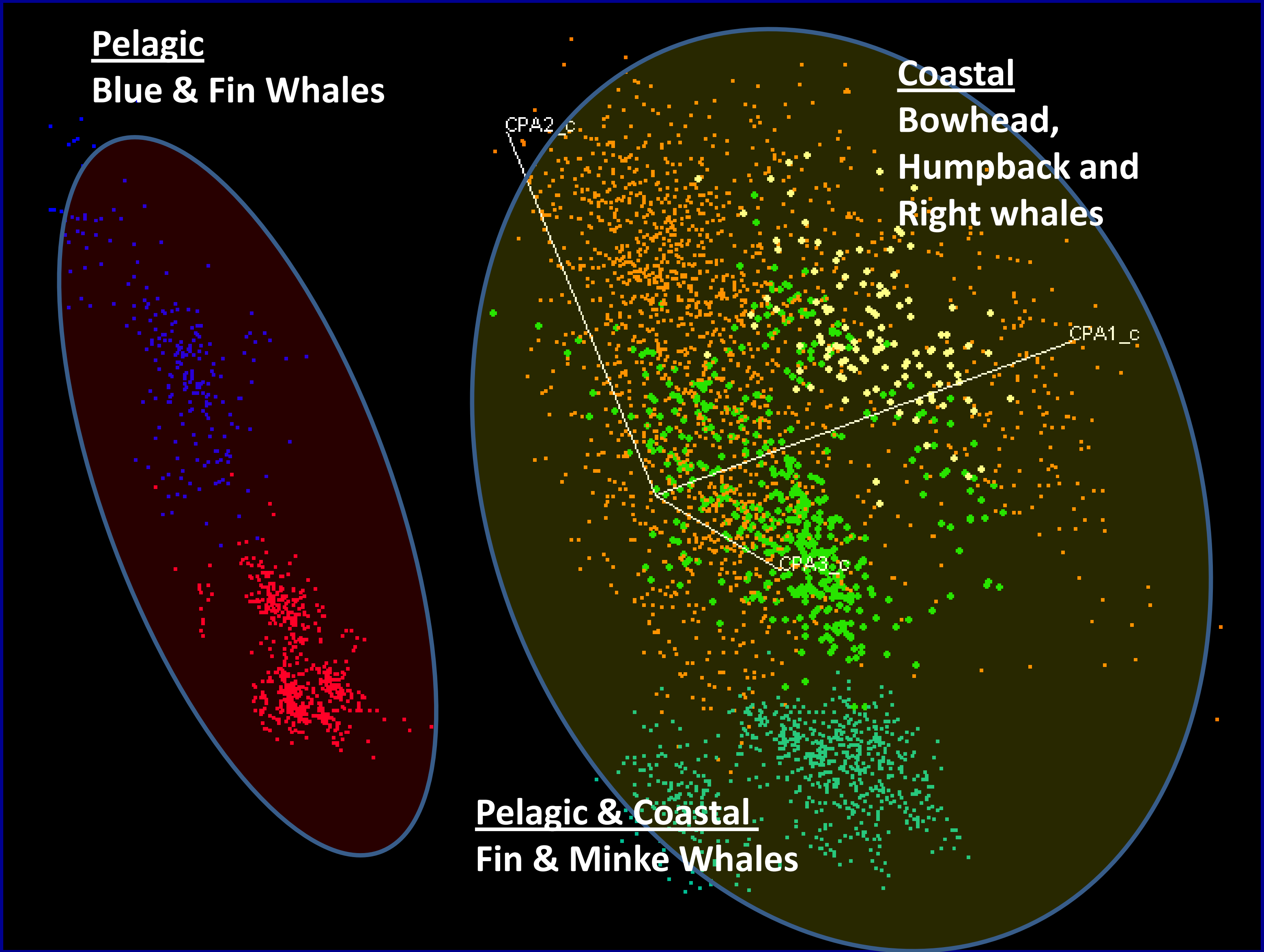
Fin

Humpback

Minke

Right

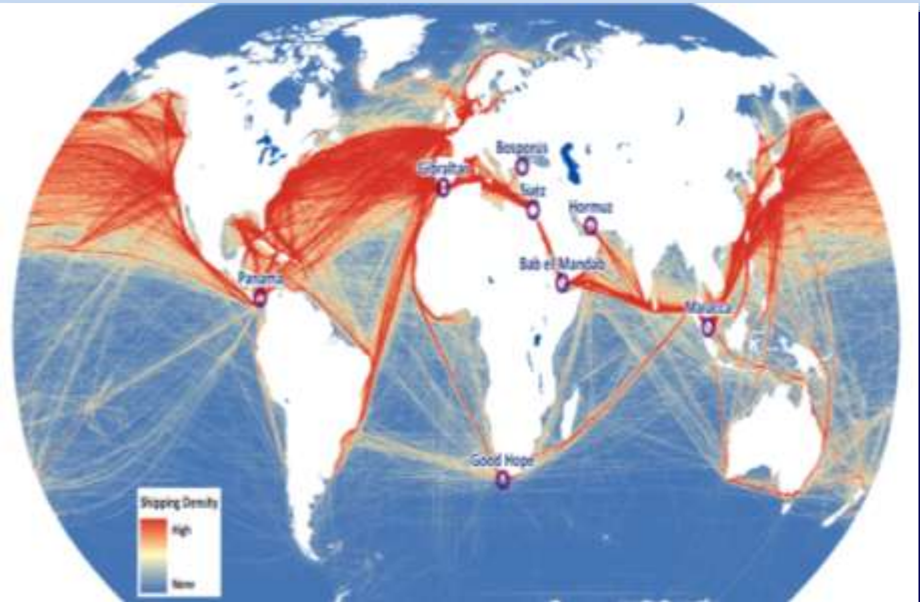
Example Feature Space for Great Whale Signals



Human activities impose huge risks to marine life over very large spatial and temporal scales.

Energy

Economics



Commercial Shipping Noise

96 % of the World's Commerce Travels on Ships, which produce high levels of low-frequency noise.

We can track ship traffic

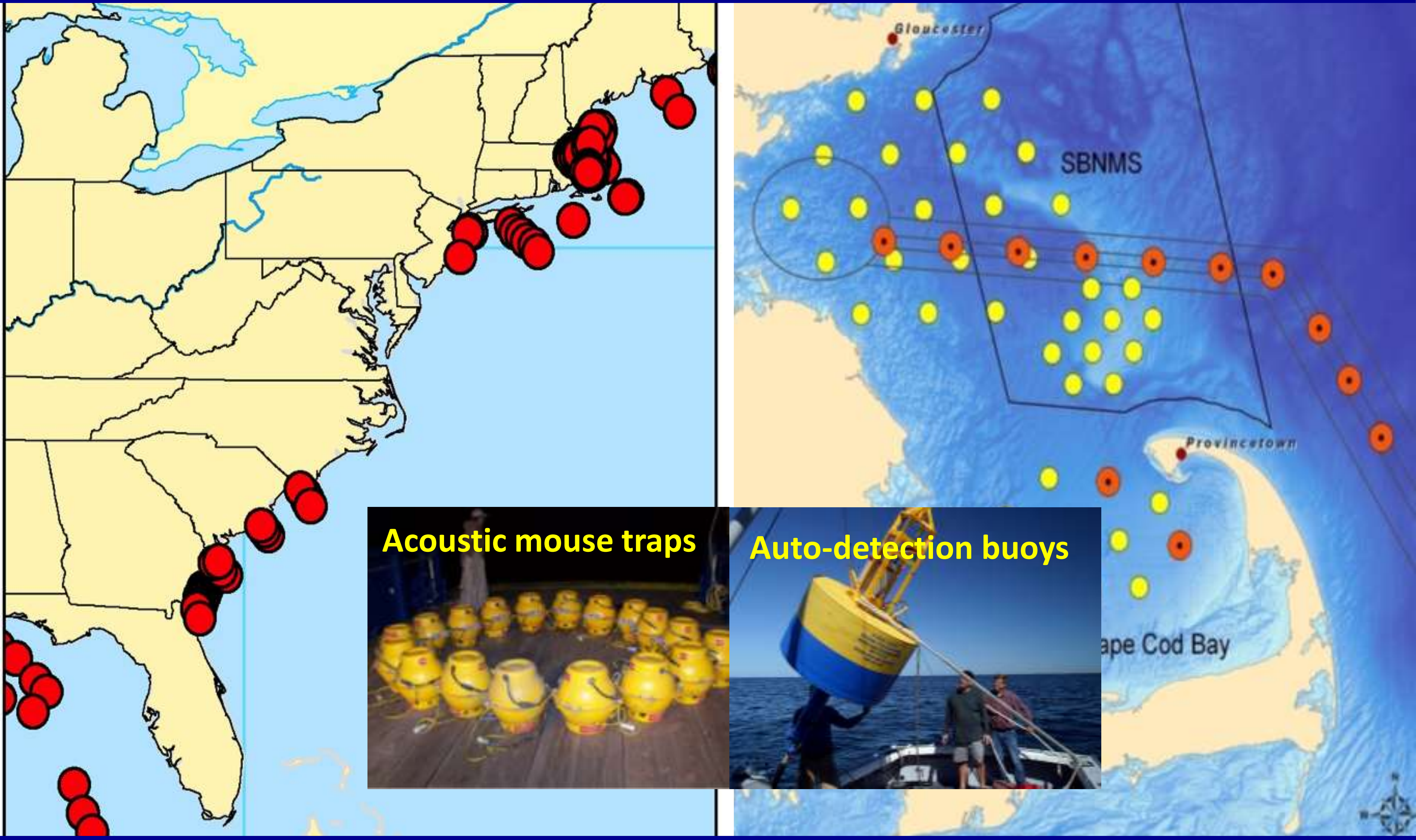
e.g. 2 months off Boston, USA

1-year Atlantic Ocean

Slide: NOAA Sound Mapping Group

Slide courtesy Stellwagen Bank National Marine Sanctuary

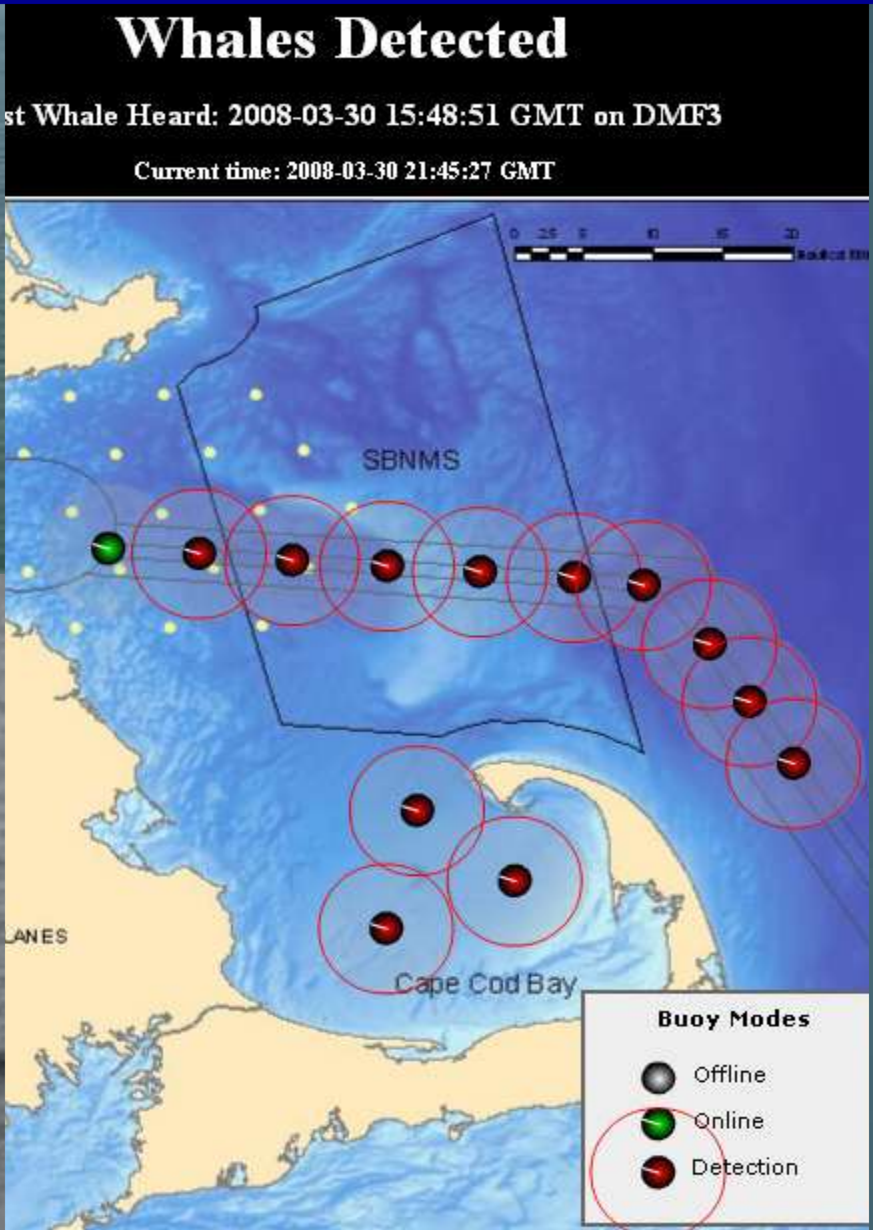
We now collect enormous amounts of acoustic data e.g. ≈ 150 years of data per year



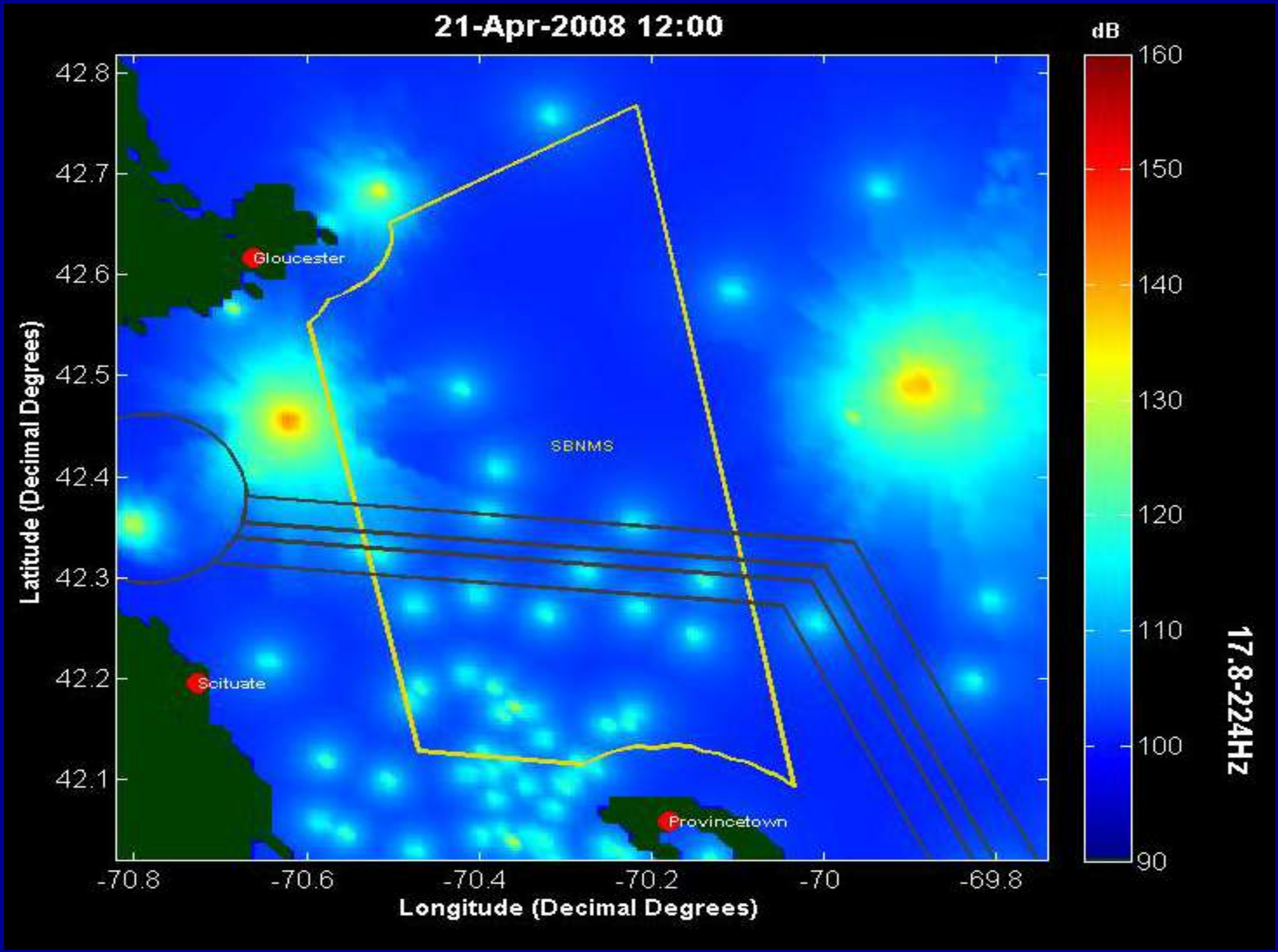
More information for Real-time monitoring can be found at: “http://www.listenforwhales.org” ©Christopher Clark

Commerce vs. Endangered Habitats

NARW-AB-Network: The First Operational Acoustic Observation System



We are beginning to translate scientific results into risk Example: endangered right whales off Boston.

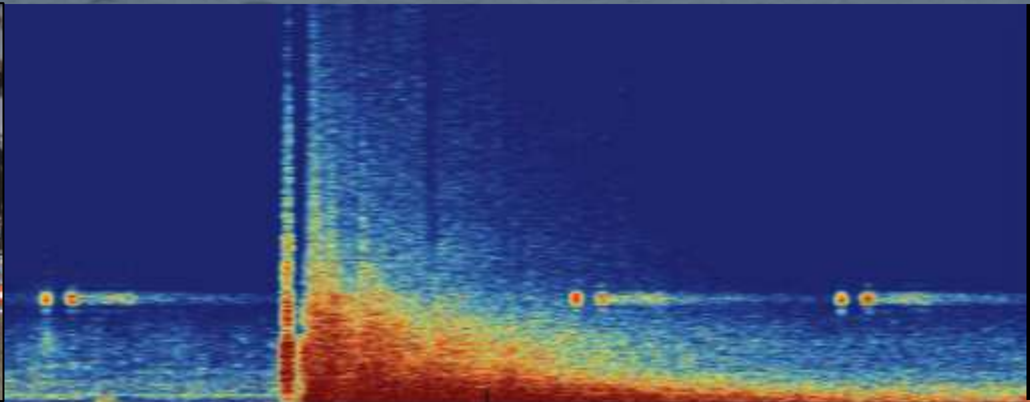


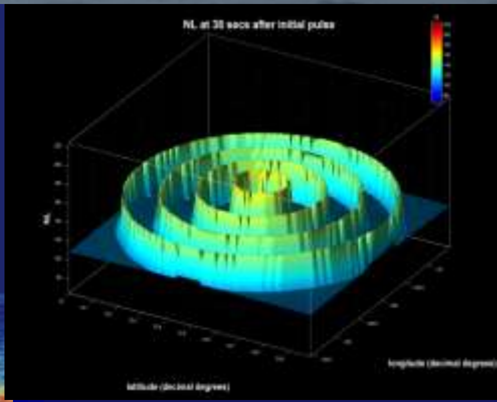
Results = Clark et al. 2009, Ellison et al. 2012, Morano et al. 2012, Hatch et al. 2012

Blue Whale Communication: pre-shipping

Blue Whale Communication – now

The scales of Seismic Airgun Surveys for hydrocarbons
Very High Noise Levels, Very Large Areas, Very Long Times





East coast: More than 300,000 seismic survey miles proposed

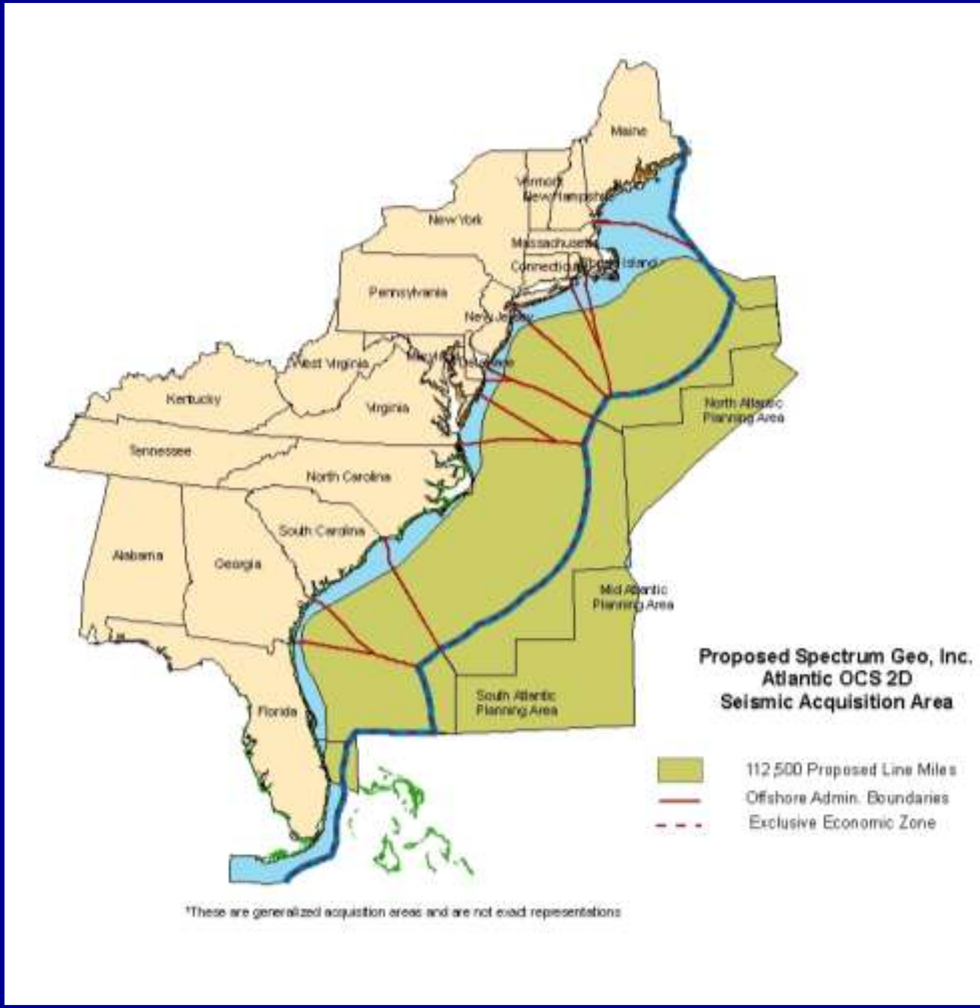
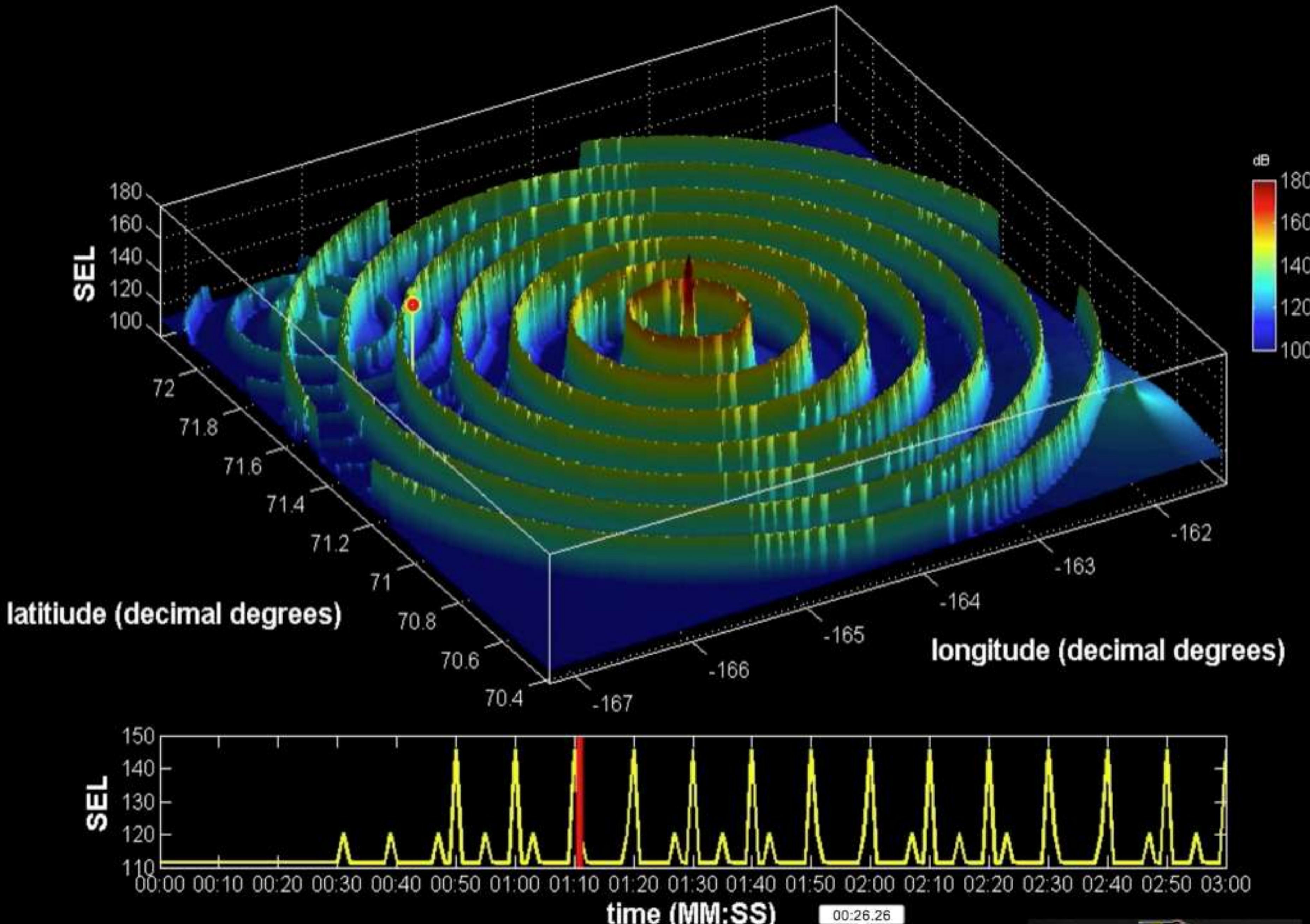


Figure from one of nine proposals submitted to Bureau of Ocean Energy & Management (BOEM) shortly after their publication of a notice-of-intent to prepare an environmental impact statement (EIS) for geophysical exploration in the Atlantic region.

SEL at 71 secs after initial pulse



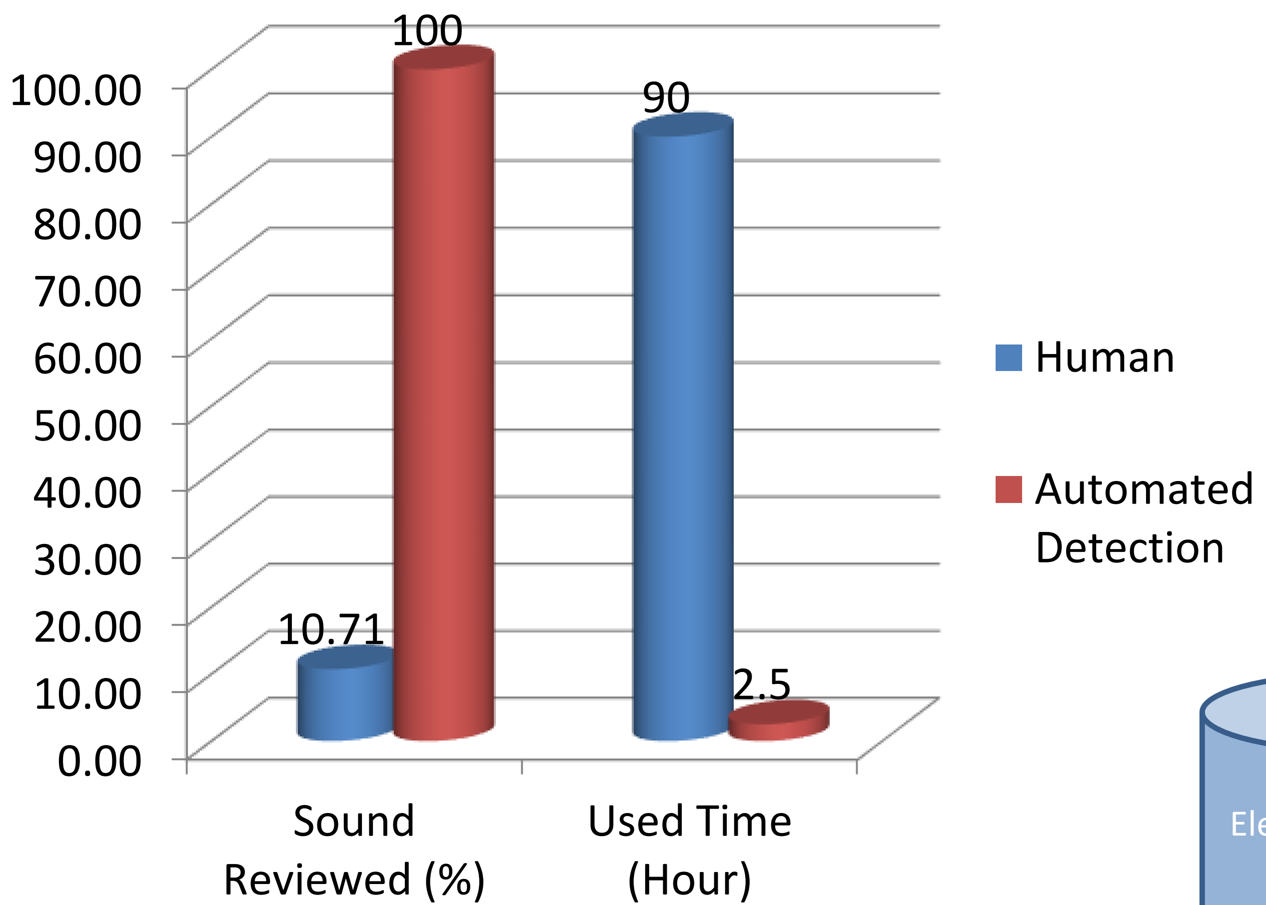
How do we process the data at appropriate

Time Scales ?

Spatial Scales ?

Frequency Scales ?

MISS TOO MUCH, TAKE TOO LONG



NOPP Grant – Advanced Detection Classification, (2012-2015)

National Oceanographic Partnership Program

Grant POP 3 year, $1M

Goal: Perform basic and applied research-development for advancing detection, classification and localization for marine bioacoustics.

Derived Requirements

Algorithm Accuracy	**Multi-year, seasonal level, hands free.**
Processing Scale	**64 – 128 nodes, multi-core (GPU later)**
Access	**Access to algorithms in ML community**
COTS	**Commercial off the shelf tools**
Client-Server	**1-2 users, focus on data products**
Processing Model	**Parallel or “tight” distributed model**

Dugan, Pourhomayoun, Shiu, Paradis, and Clark

Performance

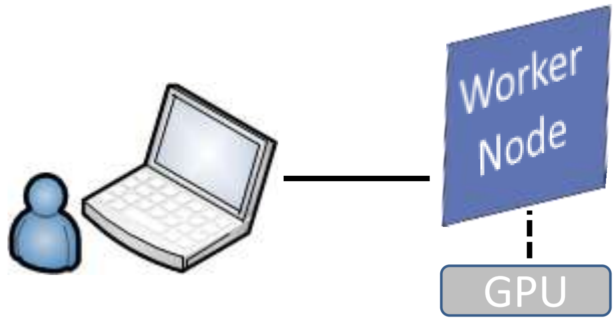
Dell Desktop versus HPC-ADA

Benchmark Detection Runtime Performance

Number Cores	Processor	Elapsed time Hrs:MM:Sec	Number of Detected and Classified Events
Dell Desktop Work Station			
1	Intel Xeon X5482 @ 3.2 GHz	2:55:12	16863
4	Intel Xeon X5482 @ 3.2 GHz	1:28:28	16863
HPC-ADA			
4	Intel Xeon X5650 @ 2.67 GHz	0:44:48	16863
12	Intel Xeon X5650 @ 2.67 GHz	0:23:24	16863
22 (10 Virtual)	Intel Xeon X5650 @ 2.67 GHz	0:19:12	16863
64	Intel Xeon X5650 @ 2.67 GHz	< 0:5:0	16863

Data Size = 2 GB; Sound Size = 114Hours, 10 Minutes, 48 Seconds (Continuous Recording); Sample Rate (Fs) = 2000 Hz

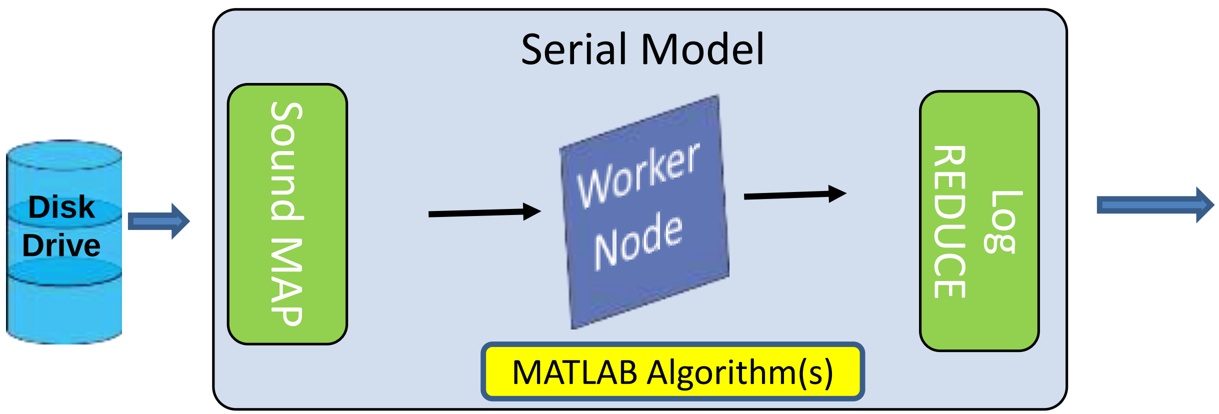
HPC Processing - Serial



Data Objects
(i.e. NARW)

+

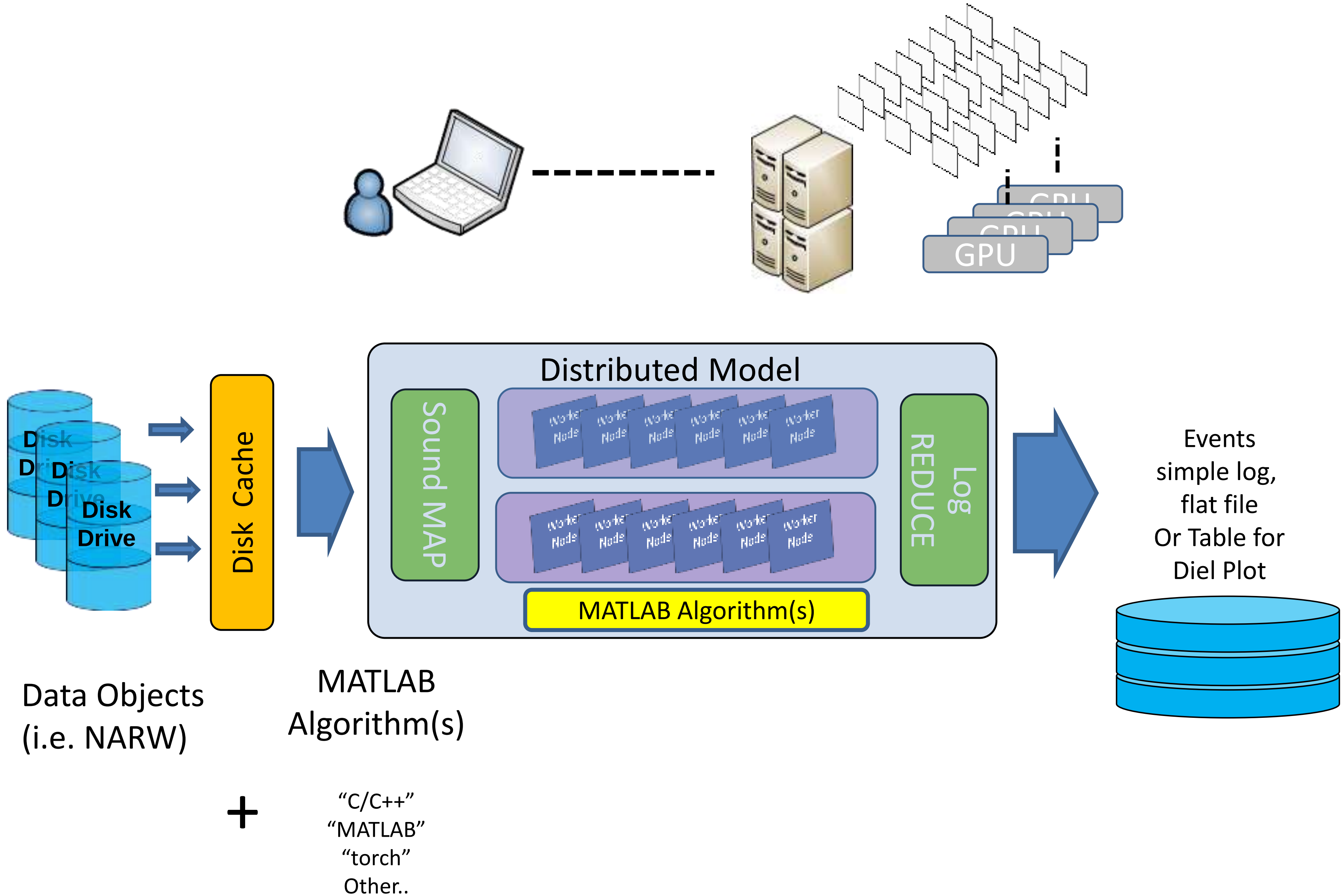
MATLAB
Algorithm(s)



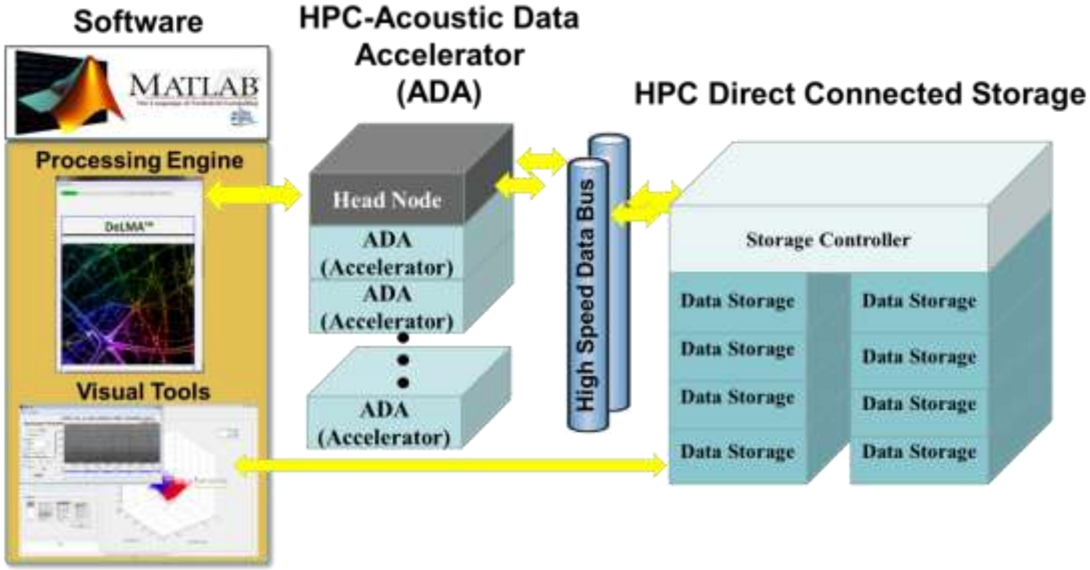
Detection
Events
“simple log
or
flat file”

“C/C++”
“MATLAB”
“torch”
Other..

HPC Processing - Distributed

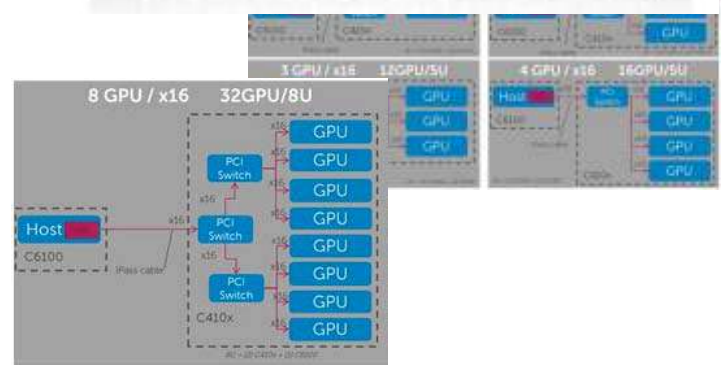


DeLMA HPC – Acoustic Data Accelerator



GPU C410x expansion

GPU C410x expansion



Specifications

- C6220 Class, Cloud Server.
- 64 Distributed Nodes, 4 mother boards.
- 192 GB RAM.
- dual Intel® Xeon® E5-2600.
- GPU support, external C410x Rack Server.
- 16 GPU's via dynamic allocation.
- Tesla NVIDIA M2075/M2090 GPUs.
- 18TB NAS with Open Indian, running NAPPit.
- Mirror fast CACHE, SDD drives.

Pulse Train Project

- Goal: Detect Minke Whale Song in Large Datasets

Example Detection Model

Spectral and Temporal Translations,

Minke

Minke

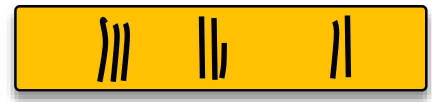
At Least Two Ranges of Temporal Resolution

Up Sweep and FM Modulated

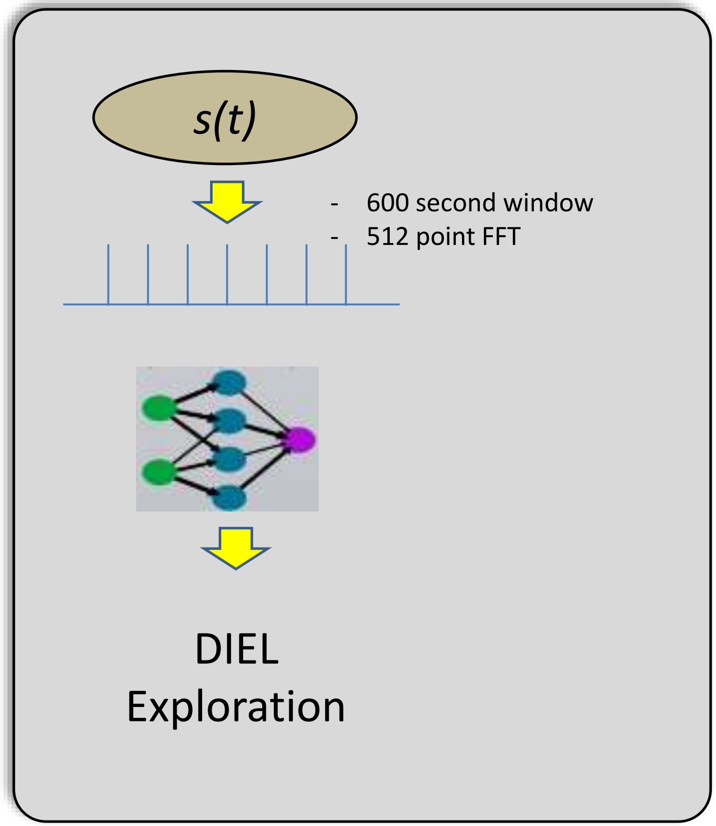
Pulse Train

Many "minke-like" Shapes at a Single Resolution

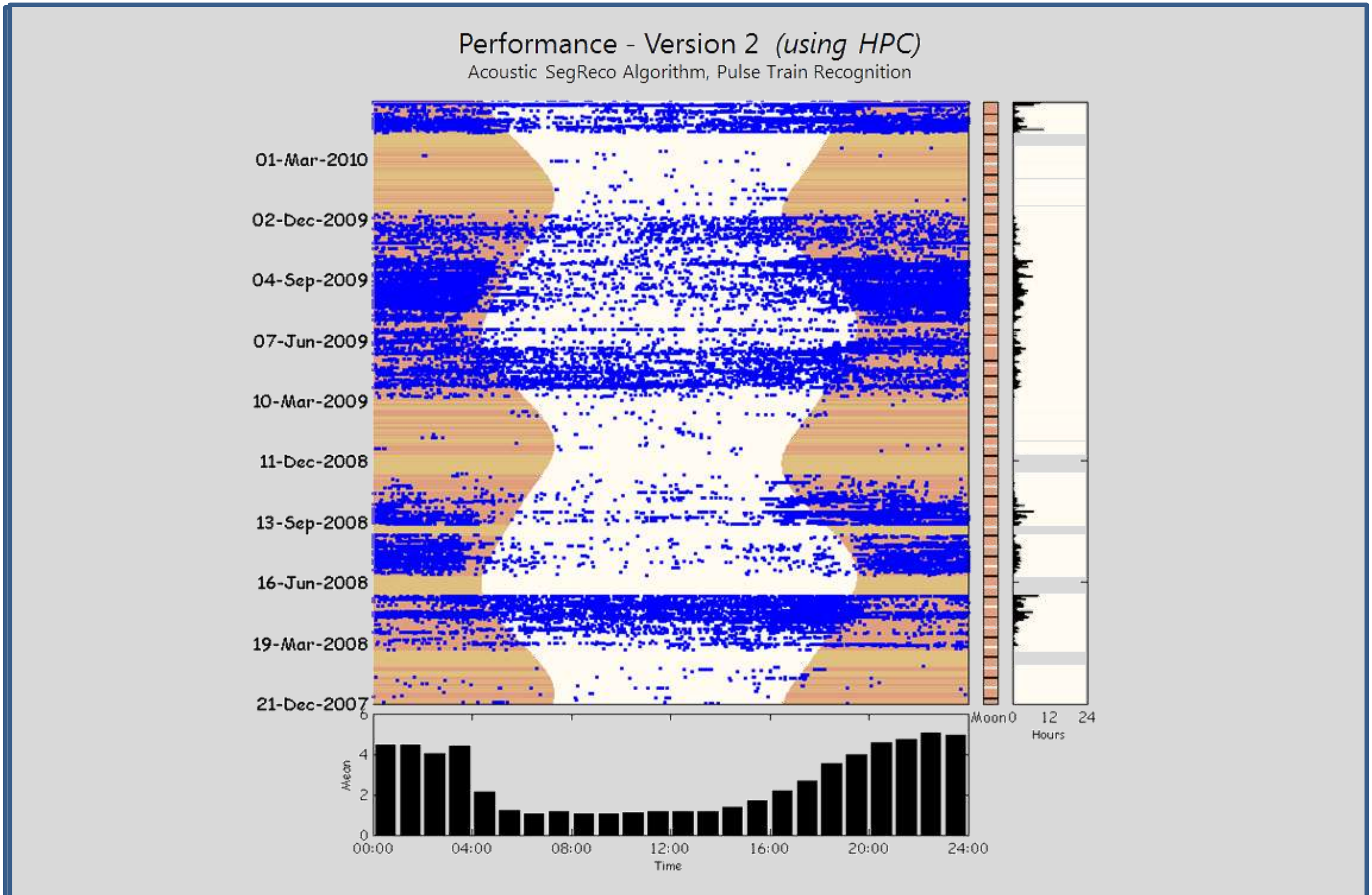
Minke



Haddock



Pulse Train Performance



Pulse Train

Right Whale

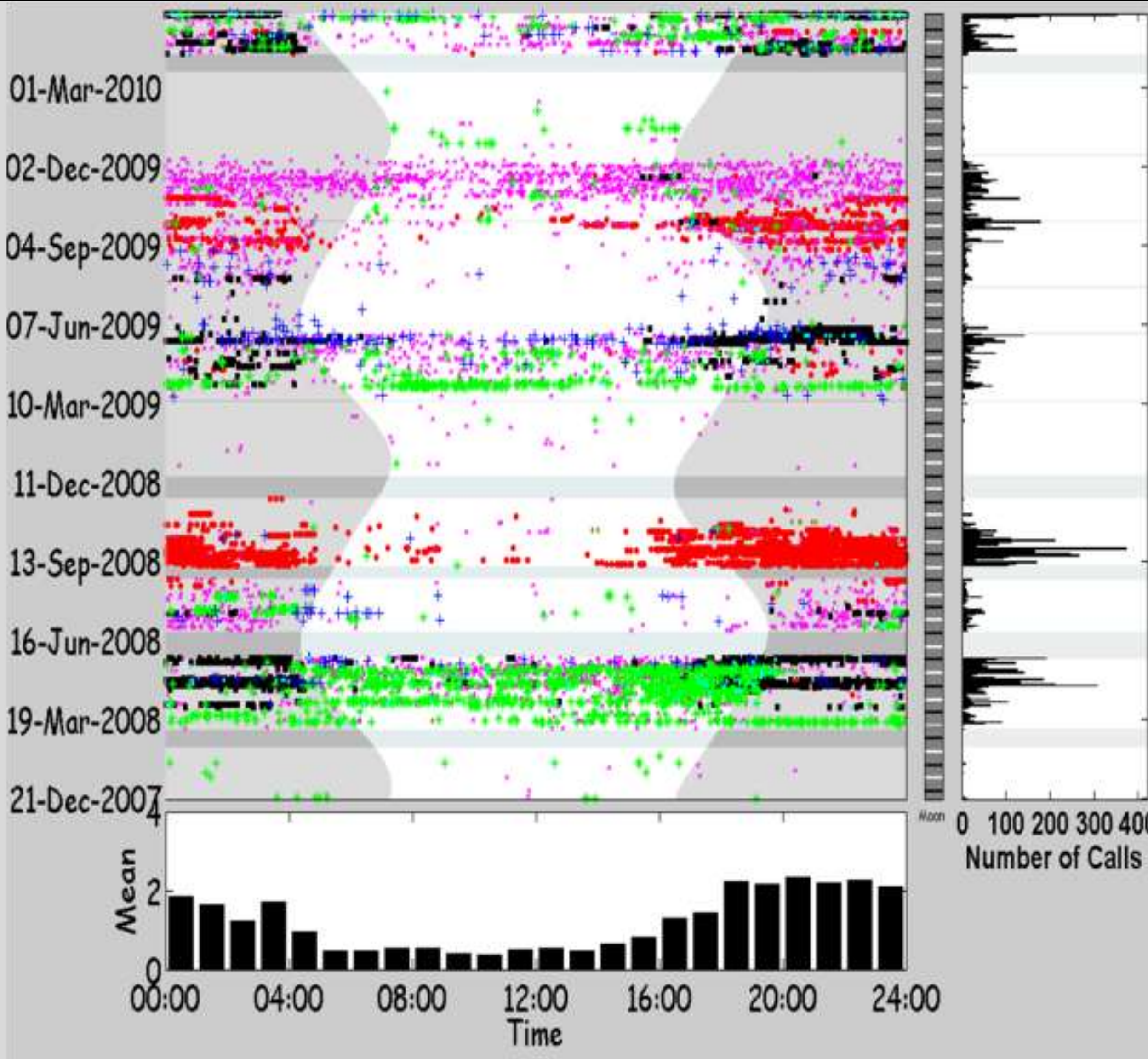
Haddock

Humpback (song)

Humpback (moan)

Unknown (PT)

Minke



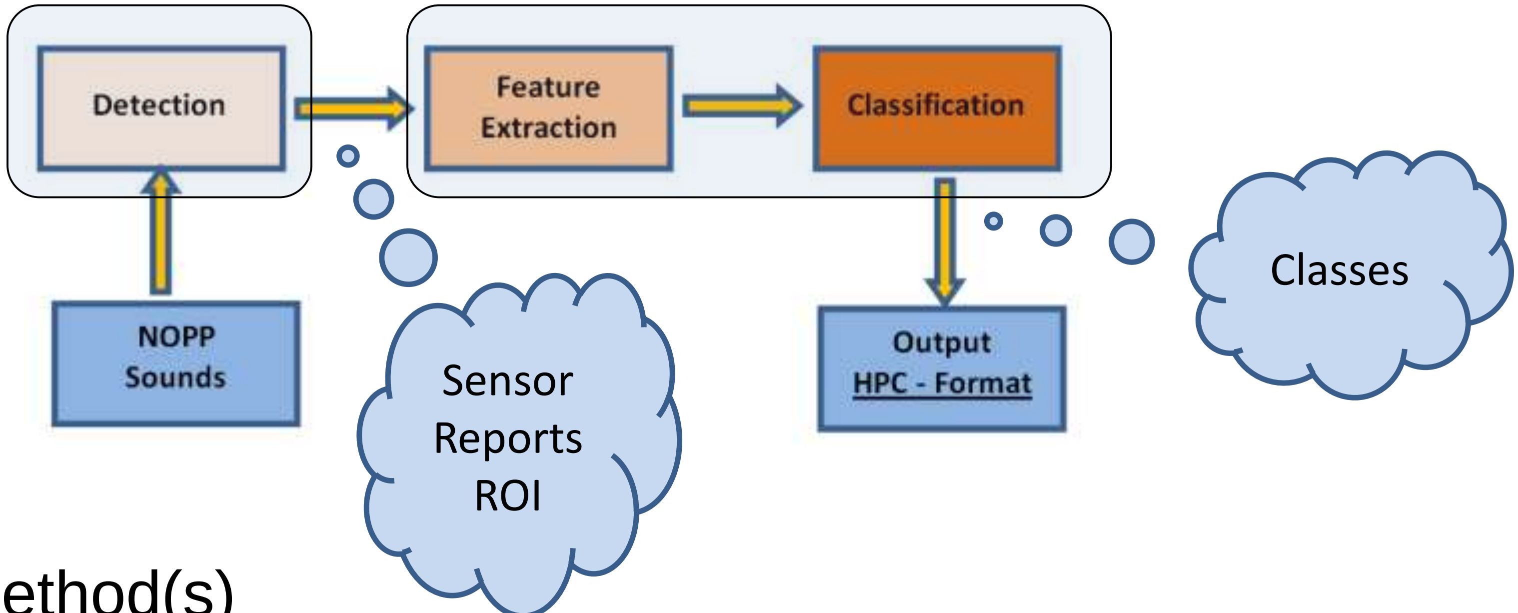
NRW Project

- Goal: Detect NARW Whale Song in Large Datasets

Take advantage of the “state of the art”.

NRW Project

Applied Segmentation Recognition (ASR)



Cornell Method(s)

1. Feature Vector Testing Model (isRAT)

   (I.Urazghildiiev, Cornell University)

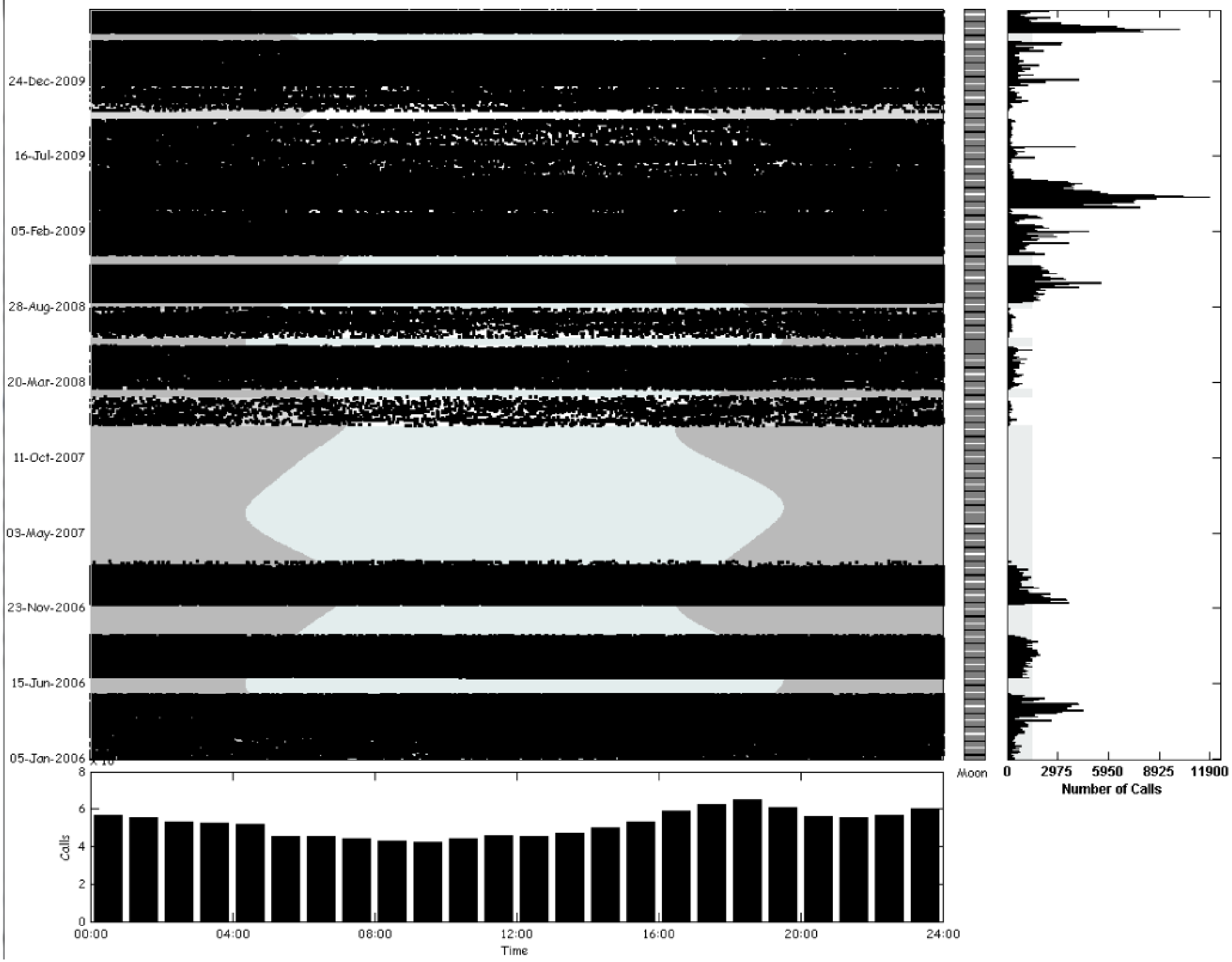
2. Connected Region Analysis (CRA)

   (M. Pourhomayoun, Cornell University)

3. Histogram Oriented Gradients (HOG)

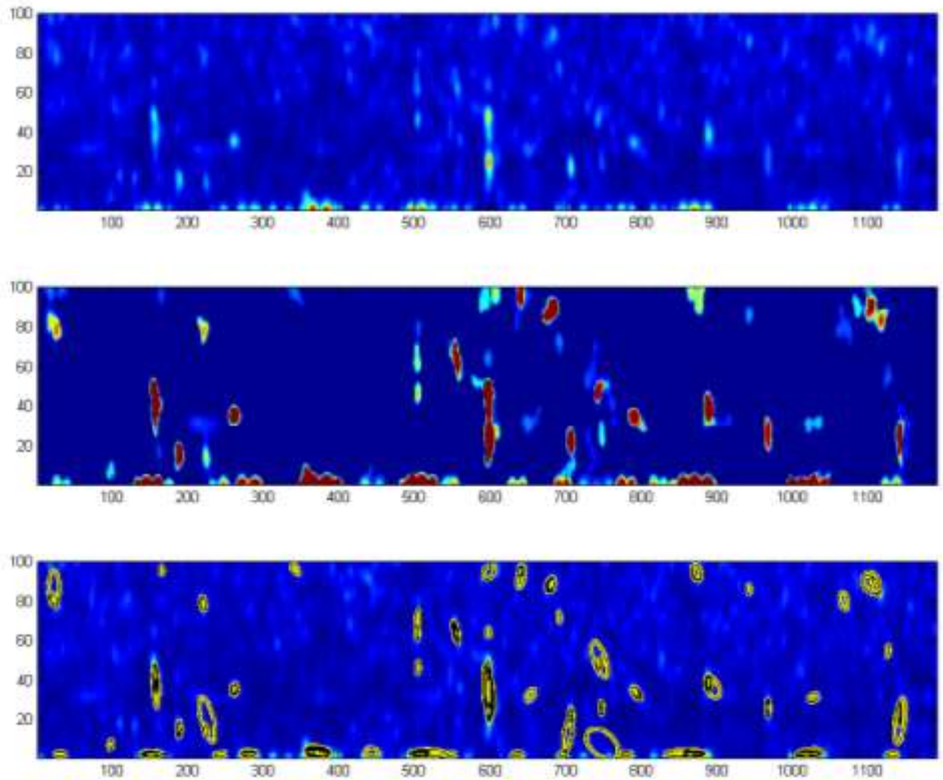
   (Y.Shiu, Cornell University)

NARW pre April 2013



Dugan, Clark, LeCun, Parijs, Shiu, Popescu, Pourhomayoun, Ponirakis and Rice

MSER Overview
Maximally Stable Extremal Regions (MSER)[1]



Sound Samples

↓

Short-time Fourier Transform

↓

MSER: find stable connected regions

↓

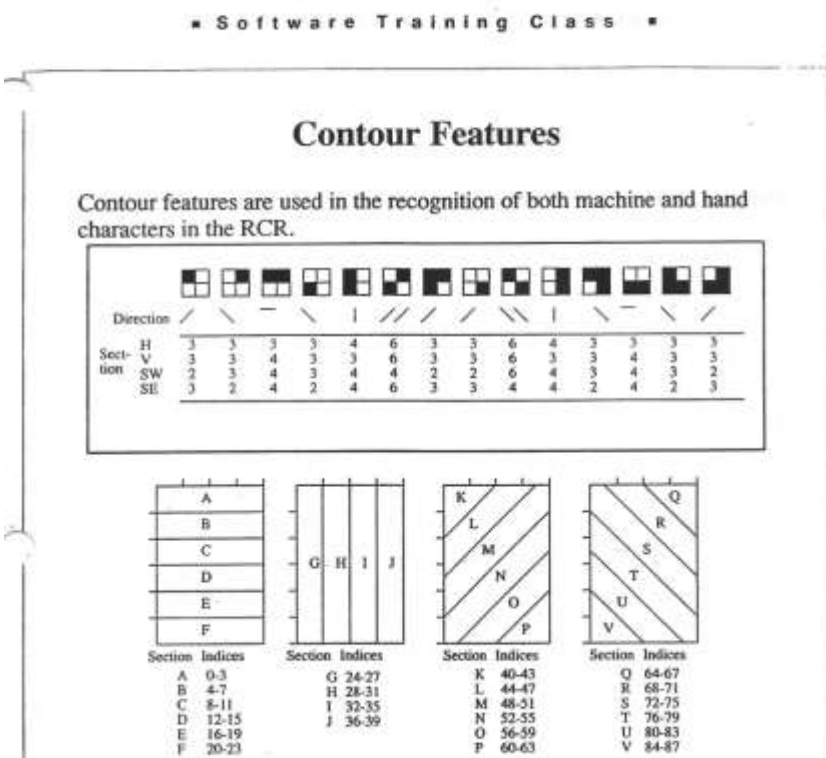
Adjacent regions are grouped into objects

↓

(t, f) of objects' center

1: Matas et al. 2002, "Robust wide baseline stereo from maximally stable extremal regions"

CRA Overview



Contour Features

Contour features are used in the recognition of both machine and hand characters in the RCR.

Direction		/	\	—	\	|	//	/	/	\\	|	\	—	\	/
Section	H	3	3	3	3	4	6	3	3	6	4	3	3	3	3
	V	3	3	4	3	3	6	3	3	6	3	3	4	3	3
	SW	2	3	4	3	4	4	2	2	6	4	3	4	3	2
	SE	3	2	4	2	4	6	3	3	4	4	2	4	2	3

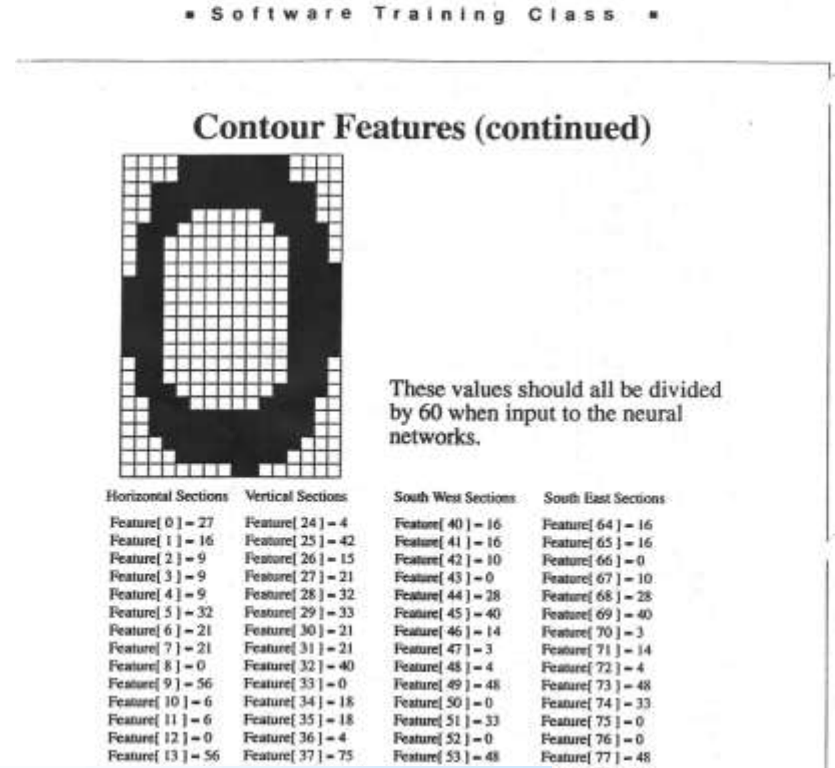
Section	Indices
A	0-3
B	4-7
C	8-11
D	12-15
E	16-19
F	20-23

Section	Indices
G	24-27
H	28-31
I	32-35
J	36-39

Section	Indices
K	40-43
L	44-47
M	48-51
N	52-55
O	56-59
P	60-63

Section	Indices
Q	64-67
R	68-71
S	72-75
T	76-79
U	80-83
V	84-87

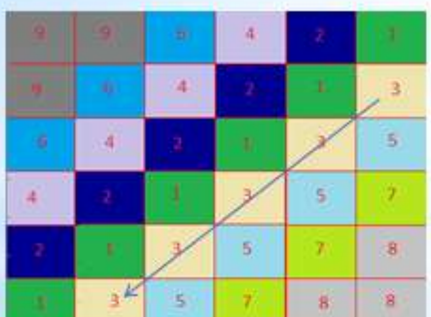
Direction	Offset
—	0
|	1
/	2
\	3



Contour Features (continued)

These values should all be divided by 60 when input to the neural networks.

Horizontal Sections	Vertical Sections	South West Sections	South East Sections
Feature[0] = 27	Feature[24] = 4	Feature[40] = 16	Feature[64] = 16
Feature[1] = 16	Feature[25] = 42	Feature[41] = 16	Feature[65] = 16
Feature[2] = 9	Feature[26] = 15	Feature[42] = 10	Feature[66] = 0
Feature[3] = 9	Feature[27] = 21	Feature[43] = 0	Feature[67] = 10
Feature[4] = 9	Feature[28] = 32	Feature[44] = 28	Feature[68] = 28
Feature[5] = 32	Feature[29] = 33	Feature[45] = 40	Feature[69] = 40
Feature[6] = 21	Feature[30] = 21	Feature[46] = 14	Feature[70] = 3
Feature[7] = 21	Feature[31] = 21	Feature[47] = 3	Feature[71] = 14
Feature[8] = 0	Feature[32] = 40	Feature[48] = 4	Feature[72] = 4
Feature[9] = 56	Feature[33] = 0	Feature[49] = 48	Feature[73] = 48
Feature[10] = 6	Feature[34] = 18	Feature[50] = 0	Feature[74] = 33
Feature[11] = 6	Feature[35] = 18	Feature[51] = 33	Feature[75] = 0
Feature[12] = 0	Feature[36] = 4	Feature[52] = 0	Feature[76] = 0
Feature[13] = 56	Feature[37] = 75	Feature[53] = 48	Feature[77] = 48
			eature[78] = 36
			eature[79] = 0
			eature[80] = 20
			eature[81] = 36
			eature[82] = 12
			eature[83] = 16
			eature[84] = 12
			eature[85] = 12
			eature[86] = 0
			eature[87] = 16

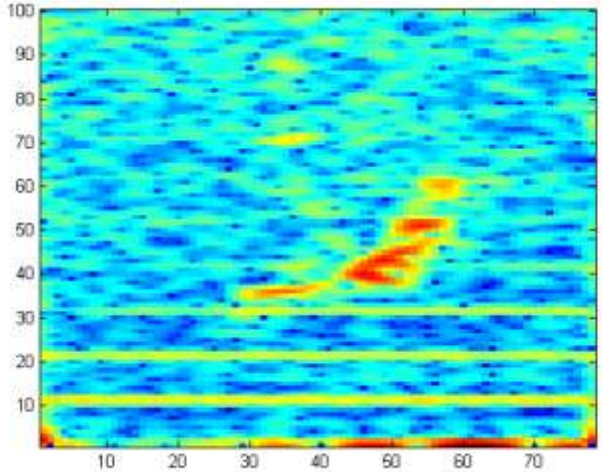


Grid Pattern used to extract the diagonal features

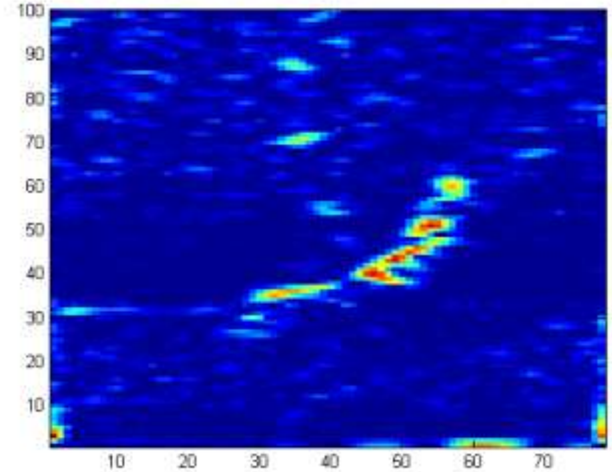
Sample Spectrogram

HOG Overview

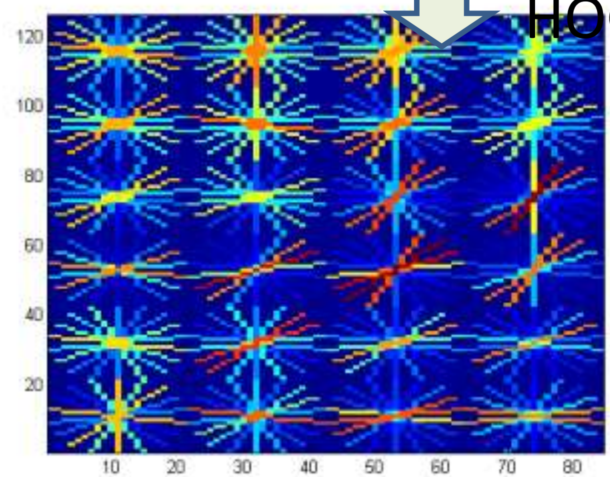
Original



Power Law



HOG



- Matas et all 2002, “Robust wide baseline stereo from ma stable extremal regions”
- Helble et al 2012, “A generalized power-law detection alg for humpback whale vocalizations”
- Dalal & Triggs 2005, “Histograms of oriented gradients fo detection”
- Vedaldi & Fulkerson 2008, “VLFeat: An Open and Portabl of Computer Vision Algorithms”

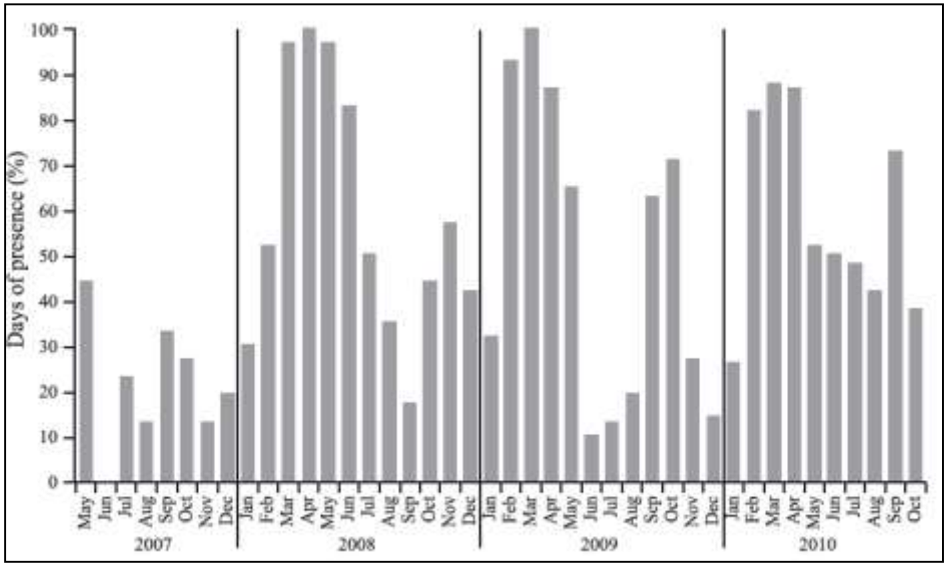
International Data Challenges – Right Whale Call
Supported by Marinexplore and Kaggle

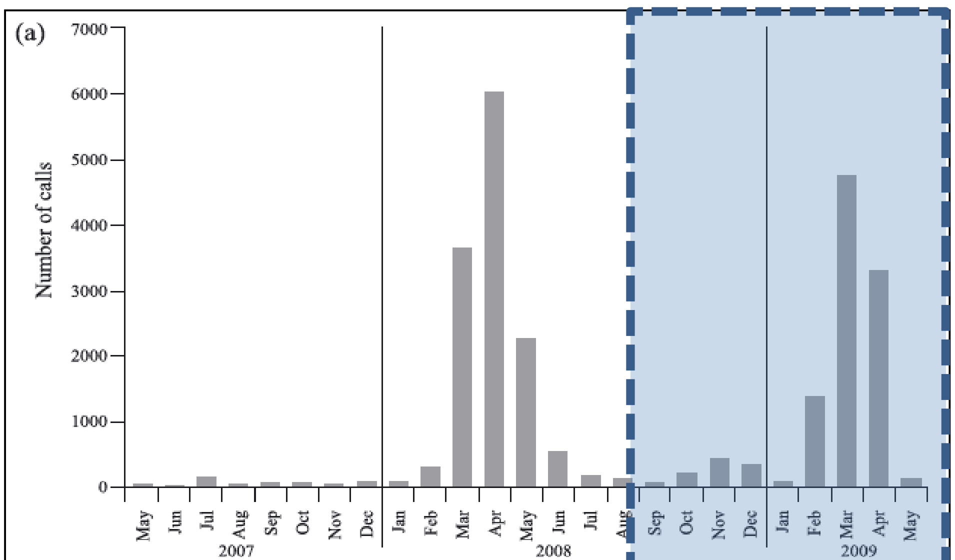
- Received over 200 entries world wide.
- Source: Auto-Buoy Data looking for NARW's.
- 70,000+ Clips: Noise, Calls.
- Problem in Classification (clip data only)

Method Name	Approach	Score	Who Submitted	Number of Features
Method 1	Template Matching + Gradient Boosting	0.9838	Dobson & Kridler	30
Method 2	Random Forest	0.9837	Nieto-Castanon	727
Method 4	ConvNet (CNN)	0.982	Cheung & Humphrey	--
HOG	HOG + Adaboost	0.964	Cornell -NYU	600
CRA	CRA+ANN	0.938	Cornell –NYU	22
Conv-Net	ConvNet (CNN)	0.926	Cornell - NYU	--

Cornell-NYU solutions finished first for (< 3 db SNR) and (< 0 db SNR) at DCLDE St. Andrews competitions.

Yearly Distribution

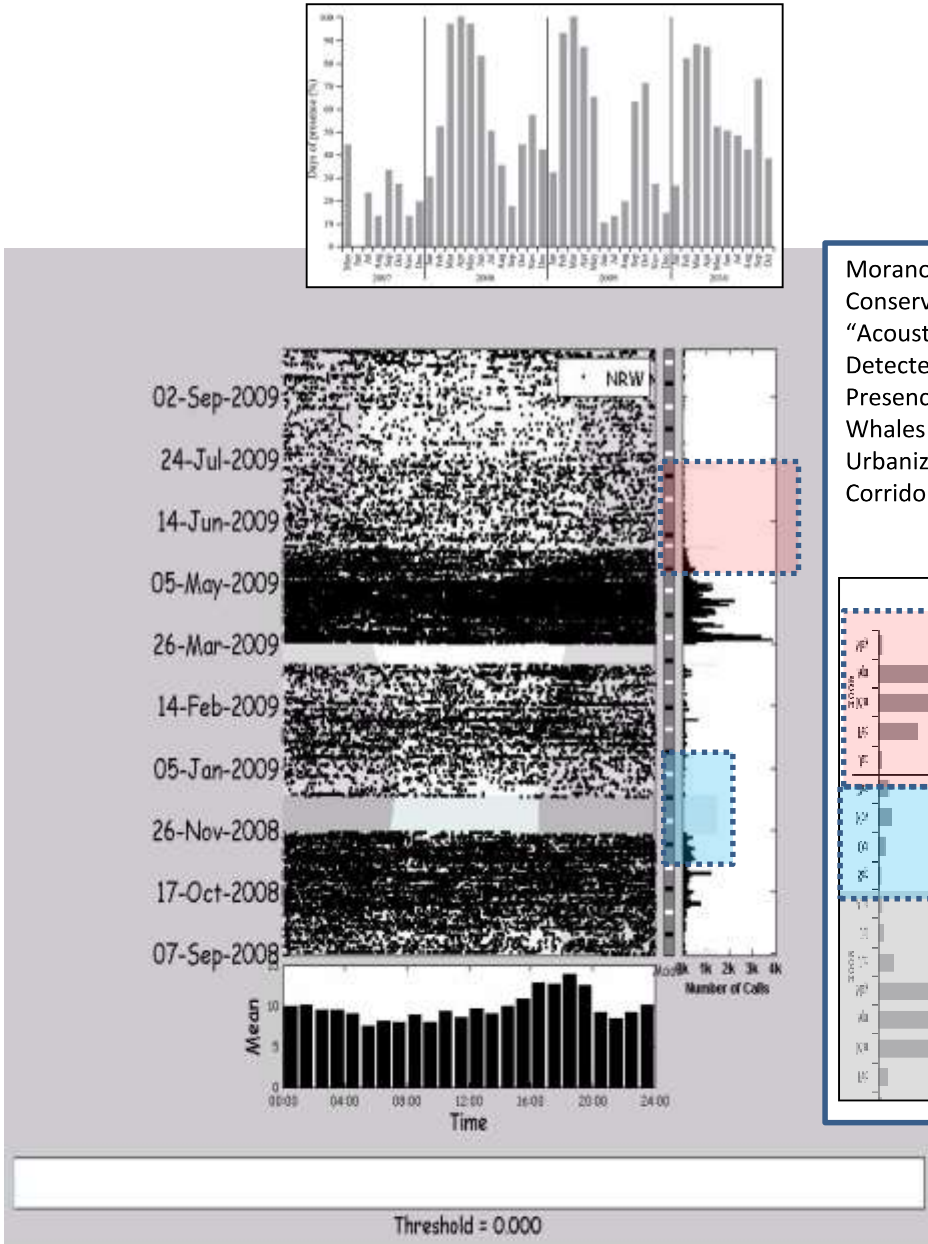




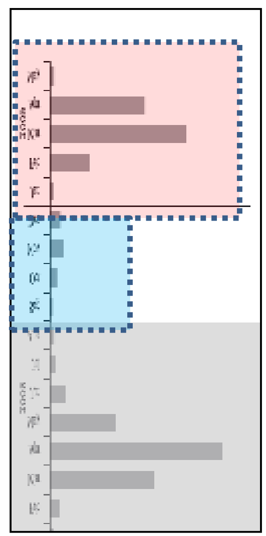
Can we find a similarity between earlier work done by biologists in the Stellwagen Sanctuary?

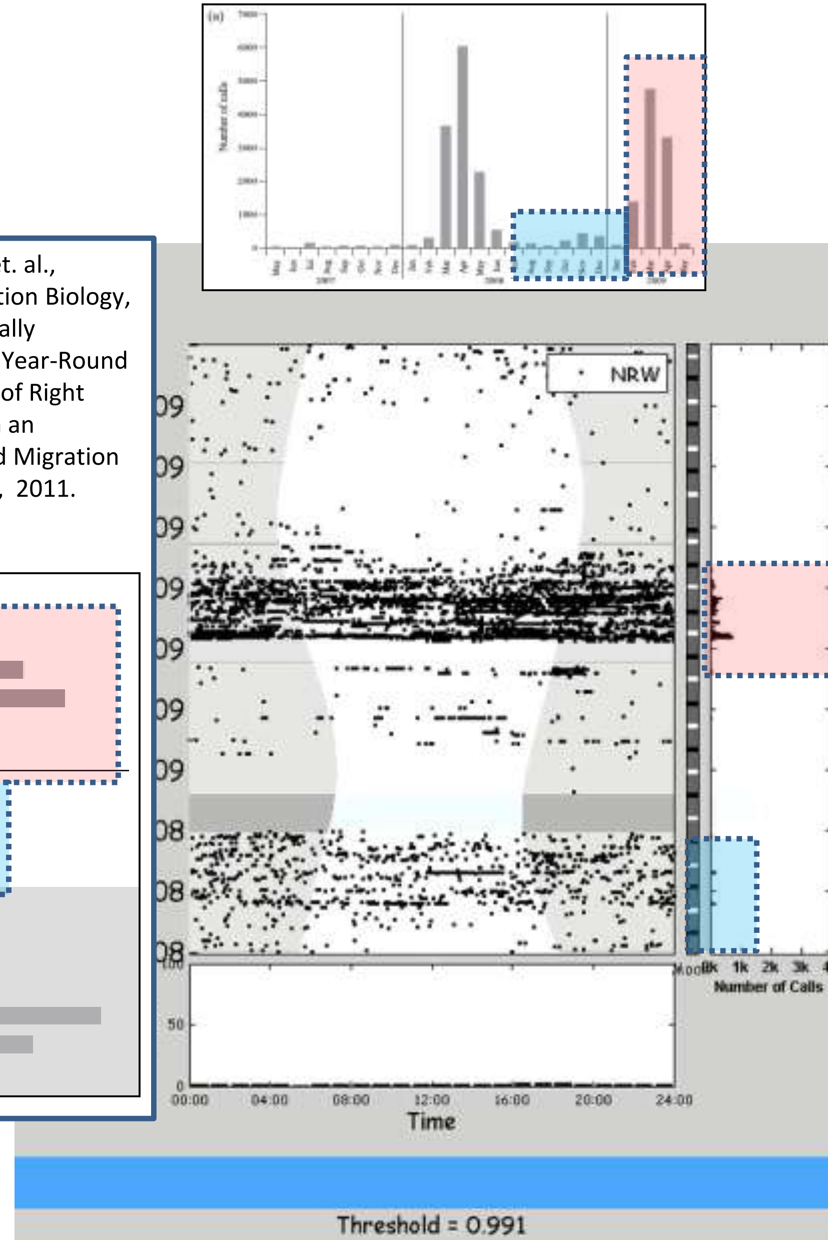
Morano et. al. (2011) measured seasonal distribution (bottom) along with animal presence (top) for the Stellwagen (NOPP) arrays. Let’s see how the algorithms work for the 2008-2009 seasonal distribution.

HOG - CRA Comparison



Morano et. al., Conservation Biology, “Acoustically Detected Year-Round Presence of Right Whales in an Urbanized Migration Corridor”, 2011.





BRP – MATLAB Team

MathWorks | *Accelerating the pace of engineering and science*

United States | Contact Us | Store

Create Account | Log In

Products & Services | Solutions | Academia | Support | User Community | Events | Company

User Stories

Share

- User Stories
- Read Stories
  - By Product
  - By Company
  - By Industry
  - By Application
  - In Academia
- Watch Videos

Cornell Bioacoustics Scientists Develop a High-Performance Computing Platform for Analyzing Big Data

View PDF

"High-performance computing with MATLAB enables us to process previously unanalyzed big data. We translate what we learn into an understanding of how human activities affect the health of ecosystems to inform responsible decisions about what humans do in the ocean and on land."

Dr. Christopher Clark, Cornell University

An acoustic analysis device used by the Bioacoustics Research Program to collect data from large baleen whales and other marine mammals. Photo courtesy Dimitri Ponirakis.

Challenge

Detect and classify animal sounds in huge sets of acoustic data acquired from oceans, fields, forests, and jungles

Solution

Develop a high-performance computing platform for acoustic data analysis using MATLAB, Parallel Computing Toolbox, and MATLAB Distributed Computing Server

Results

Special Thanks

New York University: Ross Goroshin (NYU) for support the DCL research. Xanadu Halkias for supporting ideas on integrating methods for analysis. Cornell University : Ashakur Rahaman for providing human labels for the NOPP datasets. Special thanks to Sara Keen for her support on the software and Dr. John Zollweg for integrating kaggle results. Authors would like to thank the folks from Kaggle.com and Marinexplore.com for their generous support, especially Will Cukierski for hosting Cornell datasets along with André Karpištšenko from Marinexplore, “The Ocean’s BIG Data Platform”. Special thanks to Dr. Sofie Van Parijs and Denise Risch for their help and wisdom on various aspects for the NOPP data. Special thanks to Yann LeCun and Joan Bruna from NYU …Lastly, we would like to thank our sponsors, the Office of Naval Research (ONR) and National Fish and Wildlife Foundation for making this work possible through a grant offered from the National Oceanic Partnership Program (NOPP). Lastly, very special thanks to Douglass Gillespie for hosting the 2013 workshop and providing data results.